

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2004, 09:25:55 ; Search time 1999 Seconds  
(without alignments)  
7110.754 Million cell updates/sec

Title: US-10-020-540A-1  
Perfect score: 476  
Sequence: 1 aactatgctgacagaat.....aacctccctgcttacaacc 476

Scoring table: OLIGO NUC  
Gapop 60.0 , Gapext 60.0

Searched: 27513289 seqs, 14931090276 residues

Word size : 0  
Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : EST.\*  
1: em\_estba.\*  
2: em\_esthum.\*  
3: em\_estlin.\*  
4: em\_estm.\*  
5: em\_estov.\*  
6: em\_estpl.\*  
7: em\_estro.\*  
8: em\_hrc.\*  
9: gb\_est1.\*  
10: gb\_est2.\*  
11: gb\_hrc.\*  
12: gb\_est3.\*  
13: gb\_est4.\*  
14: gb\_est5.\*  
15: em\_estfun.\*  
16: em\_estom.\*  
17: em\_gss\_hum.\*  
18: em\_gss\_inv.\*  
19: em\_gss\_pln.\*  
20: em\_gss\_vrt.\*  
21: em\_gss\_fun.\*  
22: em\_gss\_man.\*  
23: em\_gss\_mus.\*  
24: em\_gss\_pro.\*  
25: em\_gss\_rod.\*  
26: em\_gss\_phg.\*  
27: em\_gss\_vrl.\*  
28: gb\_gss1.\*  
29: gb\_gss2.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES					
Result No.	Score	Match	Length	ID	Description
1	22	4.6	279	29	CE765480
2	21	4.4	373	12	BG359460
3	21	4.4	594	28	AZ756615
4	21	4.4	595	28	AZ756578

5	21	4.4	602	28	AZ756590
6	21	4.4	772	12	BG483395
7	21	4.4	826	29	CC969844
8	20	4.2	196	10	BE949260
9	20	4.2	216	9	AA759629
10	20	4.2	247	10	AW913044
11	20	4.2	287	14	T38551
12	20	4.2	294	10	BF399080
13	20	4.2	317	10	BB747879
14	20	4.2	318	10	BF463976
15	20	4.2	331	9	AA410171
16	20	4.2	362	12	EG792711
17	20	4.2	362	12	EG817350
18	20	4.2	372	10	BB791337
19	20	4.2	385	9	AA250251
20	20	4.2	390	9	AA083503
21	20	4.2	397	13	BY638039
22	20	4.2	399	13	BY404218
23	20	4.2	400	13	BY656976
24	20	4.2	401	13	C87655
25	20	4.2	405	13	BY631080
26	20	4.2	408	13	BY646779
27	20	4.2	408	13	BY656917
28	20	4.2	409	9	AA396378
29	20	4.2	415	13	BY612244
30	20	4.2	417	10	BF147358
31	20	4.2	419	13	BY623467
32	20	4.2	420	10	BB786872
33	20	4.2	421	12	BM250201
34	20	4.2	423	9	AA112100
35	20	4.2	424	10	BB781226
36	20	4.2	424	13	BY640680
37	20	4.2	425	10	BB743084
38	20	4.2	427	9	AA983024
39	20	4.2	429	13	BY535622
40	20	4.2	442	10	BB782299
41	20	4.2	445	14	CA885627
42	20	4.2	450	10	BB796119
43	20	4.2	455	9	AA959669
44	20	4.2	457	13	BY638518
45	20	4.2	458	10	BB795665

ALIGNMENTS

RESULT 1  
CE765480  
LOCUS  
DEFINITION  
tigr-gss-dog-17000370758071 Dog Library Canis familiaris genomic,  
genomic survey sequence.  
ACCESSION  
CE765480.1 GI:37106244  
VERSION  
GSS.  
KEYWORDS  
SOURCE  
ORGANISM  
Canis familiaris (dog)  
Canis familiaris  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
REFERENCE  
1 (bases 1 to 279)  
Kirkness, E.F., Bafna, V., Halpern, A.L., Levy, S., Remington, K.,  
Rusch, D.B., Delcher, A.L., Pop, M., Wang, W., Fraser, C.M. and  
Venter, J.C.  
The dog genome: survey sequencing and comparative analysis  
Science 301 (5641), 1898-1903 (2003)  
JOURNAL  
MEDLINE  
22875432  
PUBMED  
14512627  
COMMENT  
Contact: Kirkness EF  
The Institute for Genomic Research  
Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,  
Rockville, MD 20850, USA  
Tel: 301-838-0200  
Fax: 301-838-0208  
Email: ekirknes@tigr.org

```

Class: shotgun.
Location/Qualifiers
1..279
/organism="Canis familiaris"
/mol_type="genomic DNA"
/strains="Standard Poodle"
/db_xref="taxon:9615"
/clone_lib="Dog Library"
/note="Site 1: BstXI; Libraries were prepared from
peripheral blood"

ORIGIN
Query Match 4.6%; Score 22; DB 29; Length 279;
Best Local Similarity 100.0%; Pred. No. 3.8;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 382 TATTATCTCTGCTTTTGTCT 403
Db 182 TATTATCTCTGCTTTTGTCT 203

RESULT 2
BG359460
LOCUS BG359460 373 bp mRNA linear EST 28-NOV-2001
DEFINITION sac24e07.y1 Gm-cl051 Glycine max cDNA clone GENOME SYSTEMS CLONE
ID: Gm-cl051-3278 5', mRNA sequence.
ACCESSION BG359460
VERSION BG359460.1 GI:13240151
KEYWORDS EST.
SOURCE Glycine max (soybean)
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.
REFERENCE 1 (bases 1 to 373)
AUTHORS Shoemaker,R., Keim,P., Vodkin,L., Erpelting,J., Coryell,V.,
Khanna,A., Bolla,B., Marra,M., Hillier,B., Kucaba,T., Martin,J.,
Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M.,
Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N.,
Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M.,
McGann,R., Waterston,R. and Wilson,R.
Public Soybean-EST-Project
Unpublished (1999)
CONTACT: Shoemaker,R./Public Soybean EST Project
Public Soybean EST-Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available through: ResGen, Invitrogen Corp. 2130
South Memorial Parkway Huntsville, AL 35801 For further information
call: (800)-533-4363 or contact via email: ccu@resgen.com
High quality sequence stop: 353.

FEATURES
source
1..373
/organism="Glycine max"
/mol_type="mRNA"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-cl051-3278"
/tissue_type="floral meristematic mRNA"
/lab_host="DH10B"
/clone_lib="Gm-cl051"
/note="Vector: pBluescript II SK+; Site 1: EcoRI; Site 2:
XhoI; The cDNA library was constructed from floral
meristematic mRNA provided by Dr. Halina Knap of Clemson
University. Complementary DNA was synthesized from mRNA
using a primer consisting of a poly(dT) sequence with a
XhoI restriction site. EcoRI adapters were ligated to the
blunt-ended cDNA fragments followed by XhoI digestion. The
cDNA fragments were directionally cloned into the
EcoRI-XhoI restriction site of the pBluescript vector. The

```

ligated cDNA fragments were transformed into DH10B host cells (GibcoBRL). This library was constructed in the laboratory of Dr. Randy Shoemaker."

```

ORIGIN
Query Match 4.4%; Score 21; DB 12; Length 373;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 367 ATAAATATCTTCATATATAT 387
Db 213 ATAAATATCTTCATATATAT 233

RESULT 3
AZ756615
LOCUS AZ756615 594 bp DNA linear GSS 01-MAR-2001
DEFINITION ev13hl2.r1 PAX3 CASTING Library 'ev' Homo sapiens genomic clone
ev13hl2 random, genomic survey sequence.
ACCESSION AZ756615
VERSION AZ756615.1 GI:13176068
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 594)
AUTHORS Barber,T.D., Barber,M.C., Tomescu,O., Barr,F.G., Ruben,S. and
Friedman,T.B.
TITLE Identification of Target Genes Regulated by PAX3 and PAX3--FKHR in
Embryogenesis and Alveolar Rhabdomyosarcoma
Genomics 79 (3), 278-284 (2002)
JOURNAL MEDLINE
PUBMED 11863357
COMMENT Contact: Friedman TB
Laboratory of Molecular Genetics
National Institutes of Deafness and Other Communication Disorders,
National Institutes of Health
5 Research Court, Room 2A-15, Rockville, MD 20850, USA
Tel: 301 402 7580
Fax: 301 496 7882
Email: friedman@nidcd.nih.gov
Plate: 13 row: h column: 12
Seq primer: -28M13 reverse primer (Amersham)
Class: random plasmid subclone.
Location/Qualifiers
1..594
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="ev13hl2"
/sex="Male"
/lab_host="DH10B"
/clone_lib="PAX3 CASTING Library 'ev'"
/note="Vector: pGEM-T Easy; Human genomic DNA was
partially digested with Sau3AI, ligated to ds linkers,
and enriched for binding to human PAX3Q+ protein using a
Whole Genome PCR-based strategy. DNA fragments containing
putative PAX3Q+ binding sites were amplified by PCR and
cloned into pGEM-T Easy (Promega). The ligation products
were transformed into DH10B electrocompetent cells (Life
Technologies)."
```

ligated cDNA fragments were transformed into DH10B host cells (GibcoBRL). This library was constructed in the laboratory of Dr. Randy Shoemaker."

```

ORIGIN
Query Match 4.4%; Score 21; DB 28; Length 594;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 296 AGATAAGACTGACCCAGGCC 316
Db 189 AGATAAGACTGACCCAGGCC 209
```

```

RESULT 4
AZ756578      .595 bp   DNA      linear   GSS 01-MAR-2001
LOCUS          ev13e08.r1 PAX3 CASTING Library 'ev' Homo sapiens genomic clone
DEFINITION     ev13e08 random, genomic survey sequence.
ACCESSION      AZ756578
VERSION        AZ756578.1 GI:13176030
KEYWORDS       GSS.
SOURCE         Homo sapiens (human)
ORGANISM       Homo sapiens
               Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 595)
AUTHORS        Barber,T.D., Barber,M.C., Tomescu,O., Barr,F.G., Ruben,S. and
               Friedman,T.B.
TITLE          Identification of Target Genes Regulated by PAX3 and PAX3--FKHR in
               Embryogenesis and Alveolar Rhabdomyosarcoma
JOURNAL         Genomics 79 (3), 278-284 (2002)
MEDLINE        21853298
PUBMED         11863357
COMMENT        Contact: Friedman TB
               Laboratory of Molecular Genetics
               National Institute on Deafness and Other Communication Disorders,
               National Institutes of Health
               5 Research Court, Room 2A-15, Rockville, MD 20850, USA
               Tel: 301 402 7580
               Fax: 301 496 7882
               Email: friedman@nidcd.nih.gov
               Plate: 13 row: e column: 08
               Seq primer: -28M13 reverse primer (Amersham)
               Class: random plasmid subclone.
               Location/Qualifiers
               1..595
               /organism="Homo sapiens"
               /mol_type="genomic DNA"
               /db_xref="taxon:9606"
               /clone="ev13e08"
               /sex="Male"
               /lab_host="DH10B"
               /clone_lib="PAX3 CASTING Library 'ev'"
               /note="vector: pGEM-T Easy; Human genomic DNA was
               partially digested with Sau3AI, ligated to ds linkers,
               and enriched for binding to human PAX3DQ+ protein using a
               Whole Genome PCR-based strategy. DNA fragments containing
               putative PAX3DQ+ binding sites were amplified by PCR and
               cloned into pGEM-T Easy (Promega). The ligation products
               were transformed into DH10B electrocompetent cells (Life
               Technologies)."
```

```

FEATURES
    source
    1..595
    /organism="Homo sapiens"
    /mol_type="genomic DNA"
    /db_xref="taxon:9606"
    /clone="ev13e08"
    /sex="Male"
    /lab_host="DH10B"
    /clone_lib="PAX3 CASTING Library 'ev'"
    /note="vector: pGEM-T Easy; Human genomic DNA was
    partially digested with Sau3AI, ligated to ds linkers,
    and enriched for binding to human PAX3DQ+ protein using a
    Whole Genome PCR-based strategy. DNA fragments containing
    putative PAX3DQ+ binding sites were amplified by PCR and
    cloned into pGEM-T Easy (Promega). The ligation products
    were transformed into DH10B electrocompetent cells (Life
    Technologies)."
```

```

ORIGIN
Query Match      4.4%; Score 21; DB 28; Length 602;
Best Local Similarity 100.0%; Pred. NO. 16;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 296 AGATAAGACTGACCCAGGCC 316
    |||||
DB 199 AGATAAGACTGACCCAGGCC 219
    |||||

RESULT 6
BG483395
LOCUS          BG483395
DEFINITION     602504267F1 NIH_MGC_77 Homo sapiens cDNA clone IMAGE:4617617 5',
               mRNA sequence.
ACCESSION      BG483395
VERSION        BG483395.1 GI:13415674
KEYWORDS       EST.
SOURCE         Homo sapiens (human)
ORGANISM       Homo sapiens
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE      1 (bases 1 to 772)
AUTHORS        NIH-MGC http://mgc.nci.nih.gov/.
               National Institutes of Health, Mammalian Gene Collection (MGC)
               Unpublished (1999)
JOURNAL        Contact: Robert Strausberg, Ph.D.
               Email: c9apbs-r@mail.nih.gov
               Tissue Procurement: CLONTECH Laboratories, Inc.
               cDNA Library Preparation: CLONTECH Laboratories, Inc.
               cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
               DNA Sequencing by: Incyte Genomics, Inc.
               Clone distribution: MGC clone distribution information can be
               found through the I.M.A.G.E. Consortium/ILNL at:
               http://image.llnl.gov
               Plate: LCM1374 row: f column: 18
               High quality sequence stop: 658.
               Location/Qualifiers
```

```

FEATURES
    source
    1..772
    /organism="Homo sapiens"
    /mol_type="mRNA"
    /db_xref="IMAGE:4617617 5'"
    /clone="602504267F1 NIH_MGC_77"
    /clone_lib="Homo sapiens cDNA clone IMAGE:4617617 5'"
    /note="mRNA sequence.
    EST.
    Unpublished (1999)
    Contact: Robert Strausberg, Ph.D.
    Email: c9apbs-r@mail.nih.gov
    Tissue Procurement: CLONTECH Laboratories, Inc.
    cDNA Library Preparation: CLONTECH Laboratories, Inc.
    cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
    DNA Sequencing by: Incyte Genomics, Inc.
    Clone distribution: MGC clone distribution information can be
    found through the I.M.A.G.E. Consortium/ILNL at:
    http://image.llnl.gov
    Plate: LCM1374 row: f column: 18
    High quality sequence stop: 658.
    Location/Qualifiers"
```

```

source
1. .772
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4617617"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIH_MSC_77"
/Note="Organ: lung; Vector: pDNR-LIB (Clontech); Site 1: SfiI (ggcccttgcc); Site 2: SfiI (ggccattagcc); 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGCCATTAGCC-3' and 3' adaptor sequence: 5'-ATTCTAGAGCGCGCGCCGACATG-dt(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.9 kb (range 0.5-4.0 kb). 12/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA). Note: this is a NIH_MGC Library."

ORIGIN
Query Match 4.4%; Score 21; DB 12; Length 772;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 123 AAAGACGACGACGAATGGTG 143
Db 423 AAAGACGACGACGAATGGTG 443

RESULT 7
CC969844
LOCUS
DEFINITION
CC969844
ACCESSION
CC969844
VERSION
CC969844.1 GI:33825820
KEYWORDS
GSS.
SOURCE
Zea mays
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 826)
Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T.,
Resnick, A., Fraser, C.M., Budiman, M.A., Bedell, J.A., Rohlfing, T.,
Citek, R.W., Nurnberg, A., Robbins, D. and Lakey, N.
Consortium for Maize Genomics
Unpublished (2002)
Other GSSs: ZUAFU1TV
Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seg primer: TR
Class: sheared ends.

FEATURES
Location/Qualifiers
1. .826
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBPA0054L01"
/clone_lib="ZM_3.0.4.0_KB"
/Note="Vector: pBGSK-; Site 1: HincII; 3-4 kb 'unfiltered' genomic DNA library"

ORIGIN
Query Match 4.4%; Score 21; DB 29; Length 826;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 377 TTCATTATTATCTCTGCTTT 397

```

```

Db 15 TTCATTATTATCTCTGCTTT 35

RESULT 8
BE949260/c
LOCUS
DEFINITION
UI-M-BH3-ave-h-11-0-UI.s1 NIH_BMAP_M_S4 Mus musculus cDNA clone
UI-M-BH3-ave-h-11-0-UI 3', mRNA sequence.
BE949260
ACCESSION
BE949260.1 GI:10527019
KEYWORDS
EST.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 196)
AUTHORS
Bonaldo, M.F., Lennon, G. and Soares, M.B.
TITLE
Normalization and subtraction: two approaches to facilitate gene
discovery
JOURNAL
Genome Res. 6 (9), 791-806 (1996)
MEDLINE
97044477
PUBMED
8889548
COMMENT
Contact: Chin, H
National Institute of Mental Health
6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
20892-9643, USA
Tel: 301 443 1706
Fax: 301 443 9890
Email: mEST@mail.nih.gov
The sequence contained an oligo-dT track that was present in the
oligonucleotide that was used to prime the synthesis of first
strand cDNA and therefore this may represent a bonafide poly A
tail. The sequence tag present in the cDNA between the NotI site
and the oligo-dT track served to identify it as a clone from the
spinal cord tissue cDNA Library Preparation: M.B. Soares Lab Clone
distribution: Researchers may obtain BMAP cDNA clones from RESEARCH
GENETICS. It should be noted that Bento Soares is generating a
small number of additional specialized non-redundant arrays of BMAP
cDNAs whose availability will be considered under appropriate and
limited collaborative arrangements
Seq primer: M13 Forward
POLYA=yes.

FEATURES
Location/Qualifiers
1. .196
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UI-M-BH3-ave-h-11-0-UI"
/dev_stages="27-32 days"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NIH_BMAP_M_S4"
/Note="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; The
NIH_BMAP_M_S4 library is a subtracted library of a series,
ultimately derived from a mixture of individually tagged
normalized libraries from ten regions of the mouse brain
(cerebellum, brain stems, olfactory bulbs, hypothalamus,
cortex, amygdala, basal ganglia, pineal gland, striatum,
hippocampus) after a series of subtractions to reduce the
representation of cDNAs from which ESTs had already been
generated. The following serially subtracted libraries
were generated in this process: NIH_BMAP_M_S4,
NIH_BMAP_M_S3.3, NIH_BMAP_M_S3.2, NIH_BMAP_M_S3.1,
NIH_BMAP_M_S2, NIH_BMAP_M_S1. The subtracted library
(NIH_BMAP_M_S4) was constructed as follows: PCR amplified
cDNA inserts from NIH_BMAP_M_S3.3, NIH_BMAP_M_S3.2, and
NIH_BMAP_M_S3.1 clones from which 3' ESTs had been derived
was used as a driver in a hybridization with a pool of
the NIH_BMAP_M_S3.3, NIH_BMAP_M_S3.2, and NIH_BMAP_M_S3.1
libraries in the form of single-stranded circles. The
remaining single-stranded circles (subtracted library)

```

was purified by hydroxyapatite column chromatography, converted to double-stranded circles and electroporated into DH10B bacteria (lifestyle technologies) to generate the NIH BMAP M.S4 library. This procedure has been previously described (Bonaldo, Lennon and Soares, Genome Research 6:791-806, 1996)

TAG\_TISSUE=spinal-cord  
TAG\_LIB=NIH\_BMAP\_M.S4  
TAG\_SEQ=TCAG

## ORIGIN

Query Match 4.2%; Score 20; DB 10; Length 196;  
Best Local Similarity 100.0%; Pred. No. 38;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 369 AATTATGTTTCATTATTATT 388  
Db 40 AATTATGTTTCATTATTATT 21

## RESULT 9

AA759629/c  
LOCUS 216 bp mRNA linear EST 23-JAN-1998  
DEFINITION W56506.r1 Soares\_mammary\_gland\_NMLMG Mus musculus cDNA clone  
IMAGE:1247795 5', mRNA sequence.

ACCESSION AA759629

VERSION 1

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

## ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

## REFERENCE

1 (bases 1 to 216)  
Marra, M., Hillier, L., Allen, M., Bowles, M., Dietrich, N., Dubuque, T., Geisli, S., Kucaba, T., Lacy, M., Le, M., Martin, J., Morris, M., Schellenberg, K., Steptoe, M., Tan, F., Underwood, K., Moore, B., Theising, B., Wylie, T., Lennon, G., Soares, B., Willson, R. and Waterston, R.

## TITLE

The WashU-HMI Mouse EST Project

## JOURNAL

Unpublished (1996)

## COMMENT

Contact: Marra M/Mouse EST Project

WashU-HMI Mouse EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: mouseest@wustl.edu

This clone is available royalty-free through LLNL; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

MGI:661483

Seq primer: -28m13 rev2 ET from Amersham

High quality sequence stop: 192.

Location/Qualifiers

1..216

/organism="Mus musculus"

/mol\_type="mRNA"

/db\_xref="taxon:10090"

/clone="IMAGE:1247795"

/sex="female (lactating)"

/tissue type="mammary gland"

/lab host="DH10B"

/clone lib="Soares mammary gland NMLMG"

/note="Vector: p7T3D-Pac (Pharmacia) with a modified

polylinker; 1st strand cDNA was prepared from mammary

gland tissue from a lactating female, and was then primed

with a Not I - oligo(dT) primer. Double-stranded cDNA was

ligated to Eco RI adaptors (Pharmacia), digested with Not

I and cloned into the Not I and Eco RI sites of the

modified p7T3 vector. Library is normalized. Library

was constructed by Bento Soares and M. Fatima Bonaldo."

## ORIGIN

Query Match 4.2%; Score 20; DB 9; Length 216;

Best Local Similarity 100.0%; Pred. No. 39;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 369 AATTATGTTTCATTATTATT 388

Db 213 AATTATGTTTCATTATTATT 194

## RESULT 10

AW913044/c

LOCUS 247 bp mRNA linear EST 25-MAY-2000

DEFINITION uf48d01.y1 Soares\_mammary\_gland\_NMLMG Mus musculus cDNA clone

IMAGE:1514593 5', mRNA sequence.

ACCESSION AW913044

VERSION 1

KEYWORDS EST.

SOURCE Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 247)

NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished (1997)

Other ESTs: uf48d01.x1

Contact: Robert Strausberg, Ph.D.

Email: cgapbs@mail.nih.gov

This clone is available royalty-free through LLNL; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

MGI:941445

Seq primer: -40RP from Gibco

High quality sequence stop: 241.

Location/Qualifiers

1..247

/organism="Mus musculus"

/mol\_type="mRNA"

/db\_xref="taxon:10090"

/clone="IMAGE:1514593"

/sex="female (lactating)"

/tissue type="mammary gland"

/lab host="DH10B"

/clone lib="Soares mammary gland NMLMG"

/note="Vector: p7T3D-Pac (Pharmacia) with a modified

polylinker; 1st strand cDNA was prepared from mammary

gland tissue from a lactating female, and was then primed

with a Not I - oligo(dT) primer. Double-stranded cDNA was

ligated to Eco RI adaptors (Pharmacia), digested with Not

I and cloned into the Not I and Eco RI sites of the

modified p7T3 vector. Library is normalized. Library

was constructed by Bento Soares and M. Fatima Bonaldo."

## ORIGIN

Query Match 4.2%; Score 20; DB 10; Length 247;

Best Local Similarity 100.0%; Pred. No. 41;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 369 AATTATGTTTCATTATTATT 388

Db 208 AATTATGTTTCATTATTATT 189

## RESULT 11

T38551

LOCUS 287 bp mRNA linear EST 11-JAN-1995

DEFINITION T38551.054 S. cerevisiae strain X2180-1A Saccharomyces cerevisiae

cDNA 3' end, mRNA sequence.

ACCESSION T38551

VERSION 1

KEYWORDS EST.

SOURCE Saccharomyces cerevisiae (baker's yeast)

ORGANISM

Saccharomyces cerevisiae

Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

Saccharomycetales; Saccharomycetaceae; Saccharomycetes.

# REFERENCE

1 (bases 1 to 287)  
Weinstock, K.  
Saccharomycetes cerevisiae cDNAs  
Unpublished (1995)  
Contact: Weinstock, K. and Venter, J.C.  
The Institute for Genomic Research  
932 Clopper Rd, Gaithersburg, MD 20878  
Tel: 3018699056  
Fax: 3018699423  
Email: tdbinfo@tdb.tigr.org  
For clone availability please contact the TIGR Database  
(tdbinfo@tdb.tigr.org)  
Seq primer: M13-21.

# FEATURES

source  
1..287  
/organism="Saccharomycetes cerevisiae"  
/mol\_type="mRNA"  
/strain="X2180-1A"  
/db\_xref="taxon:4932"  
/clone\_lib="S. cerevisiae strain X2180-1A"  
/note="Vector: pBluescript SK-; Site\_1: EcoRI; Site\_2:  
XhoI"

# ORIGIN

Query Match 4.2%; Score 20; DB 14; Length 287;  
Best Local Similarity 100.0%; Pred. No. 42;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 416 CTTGATTTCATTGGAAGG 435  
DB 228 CTTGATTTCATTGGAAGG 247

# RESULT 12

BF399080/c  
LOCUS  
DEFINITION  
BF399080  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
MEDLINE  
PUBMED  
COMMENT

BF399080 294 bp mRNA linear EST 27-NOV-2000  
UI-R-CAL-bja-n-15-0-UI.s1 UI-R-CAL Rattus norvegicus cDNA clone  
UI-R-CAL-bja-n-15-0-UI 3', mRNA sequence.  
BF399080  
EST  
Rattus norvegicus (Norway rat)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.  
1 (bases 1 to 294)  
Ronald, M.F., Lennon, G. and Soares, M.B.  
Normalization and subtraction: two approaches to facilitate gene  
discovery  
Genome Res. 6 (9), 791-806 (1996)  
97044477  
889548  
Contact: Soares, MB  
Coordinated Laboratory for Computational Genomics  
University of Iowa  
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA  
Tel: 319 335 8250  
Fax: 319 335 9565  
Email: bento-soares@uiowa.edu  
The sequence contained an oligo-dT track that was present in the  
oligonucleotide that was used to prime the synthesis of first  
strand cDNA and therefore this may represent a bonafide poly A  
tail. The sequence tag present in the cDNA between the NotI site  
and the oligo-dT track served to identify it as a clone from the  
normalized hippocampus library cDNA Library Preparation: M.B.  
Soares Lab Clone Distribution: clones will be available through  
Research Genetics (www.resgen.com)  
Seq primer: M13 Forward  
POLYA-yes.  
Location/Qualifiers  
1..294

/organism="Rattus norvegicus"  
/mol\_type="mRNA"  
/strain="Sprague-Dawley"  
/db\_xref="taxon:10116"  
/clone="UI-R-CAL-bja-n-15-0-UI"  
/lab\_host="DH10B (Life Technologies)"  
/clone\_lib="UI-R-CAL"  
/note="Vector: pT73D-Pac (Pharmacia) with a modified  
polylinker; Site 1: Not I; Site 2: Eco RI; The UI-R-CAL  
library is a subtracted library derived from the following  
tissues: thalamus, cerebellum, hypothalamus, medulla, testis,  
pons, midbrain, cerebral cortex, corpus striatum, and  
hippocampus. For a detailed description of the  
library from which this clone was derived, please visit  
our web site at [rattest.eng.uiowa.edu](http://rattest.eng.uiowa.edu). The subtraction  
has been previously described in (Bonaldo, Lennon and  
Soares, Genome Research 6:791-806, 1996)  
TAG\_ISSUE=hippocampus  
TAG\_LIB=UI-R-CAL  
TAG\_SEQ=GATTG

# ORIGIN

Query Match 4.2%; Score 20; DB 10; Length 294;  
Best Local Similarity 100.0%; Pred. No. 43;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 323 AAAGAAGAAATGCAAGT 342  
DB 142 AAAGAAGAAATGCAAGT 123

# RESULT 13

BB747879/c  
LOCUS  
DEFINITION  
BB747879  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

BB747879 317 bp mRNA linear EST 15-OCT-2001  
CDNA clone F530208B02 3', mRNA sequence.  
BB747879  
EST  
Mus musculus (house mouse)  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
Akimura, T., Arakawa, T., Carninci, P., Furuno, M., Hanagaki, T.,  
Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Imotani, K.,  
Ishii, Y., Ito, M., Kawai, J., Kojima, Y., Konno, H., Kouda, M.,  
Matsuyama, T., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R.,  
Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K., Sakazume, N.,  
Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,  
Sogabe, Y., Suzuki, H., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,  
Tanaka, T., Tomaru, A., Toya, T., Watahiki, A., Yasunishi, A.,  
Muramatsu, M. and Hayashizaki, Y.  
RIKEN Encyclopedia of Mouse Full-length cDNAs (Akimura, T., et al.  
2001)  
Unpublished (2001)  
Contact: Yoshihide Hayashizaki  
Laboratory for Genome Exploration Research Group, RIKEN Genomic  
Sciences Center (GSC), Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)  
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
Tel: 81-45-503-9222  
Fax: 81-45-503-9216  
Email: genome-res@gs.c.riken.go.jp/  
URL: <http://genome.gsc.riken.go.jp/>  
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,  
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
Normalization and subtraction of cap-trapper-selected cDNAs to  
prepare full-length cDNA libraries for rapid discovery of new  
genes. Genome Res. 10 (10), 1617-1630 (2000)  
wagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,  
Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T.,  
Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kita, A.

and Hayashizaki, Y.  
RIKEN integrated sequence analysis (RISA) system--384-format  
sequencing pipeline with 384 multicapillary sequencer. Genome Res.  
10 (11), 1757-1771 (2000)  
Kono, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P.,  
Sugahara, Y. and Hayashizaki, Y.  
Computer-based methods for the mouse full-length cDNA  
encyclopedia: real-time sequence clustering for construction of a  
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
Please visit our web site (<http://genome.gsc.riken.go.jp>) for  
further details.  
e mouse tissues.

## FEATURES

## source

```
Location/Qualifiers
1. .317
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="F53020B02"
/sex="male"
/tissue_type="kidney"
/dev_stage="adult"
/lab_host="SOLR"
/clone_lib="RIKEN full-length enriched, adult male kidney"
/notes="Site 1: XhoI; Site 2: SstI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN. Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5',
GAGAGAGAGCGGCGCGCACTCGAGTGTGTTTTTTTTTTTNN 3'], cDNA was
prepared by using trehalose thermo-activated reverse
transcriptase and subsequently enriched for full-length by
cap-trapper. Second strand cDNA was prepared with the
primer adapter of sequence [5',
GAGAGAGAGAGGATCCAGAGCTCATTAATTAATTAACCCCCCCCCC 3'].
cDNA was cleaved with XhoI and SstI."
```

## ORIGIN

```
Query Match      4.2%; Score 20; DB 10; Length 317;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 369 AATTATGTTTCATTATTATT 388
      |||||
Db 36 AATTATGTTTCATTATTATT 17
```

## RESULT 14

```
BF463976      318 bp mRNA linear EST 04-DEC-2000
LOCUS      BP463976      318 bp mRNA linear EST 04-DEC-2000
DEFINITION      UI-M-CG0p-bnr-h-05-0-UI-s1 NIH_BMAP Ret4_S2 Mus musculus cDNA clone
ACCESSION      UI-M-CG0p-bnr-h-05-0-UI 3', mRNA sequence.
VERSION      BF463976
KEYWORDS      EST.
SOURCE      Mus musculus (house mouse)
ORGANISM      Mus musculus
```

```
REFERENCE      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 318)
Bonaldo, M.F., Lennon, G. and Soares, M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
```

```
JOURNAL      Genome Res. 6 (9), 791-806 (1996)
```

```
MEDLINE      97044477
```

```
PUBMED      889548
```

```
COMMENT      Contact: Chin, H
```

National Institute of Mental Health  
6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD  
20892-9643, USA  
Tel: 301 443 1706  
Fax: 301 443 9890

Email: [mSST@mail.nih.gov](mailto:mSST@mail.nih.gov)  
The sequence contained an oligo-dT track that was present in the  
oligonucleotide that was used to prime the synthesis of first  
strand cDNA and therefore this may represent a bonafide poly A  
tail. The sequence tag present in the cDNA between the NotI site  
and the oligo-dT track served to identify it as a clone from the  
retina tissue cDNA Library Preparation. M.B. Soares Lab Clone  
distribution: Researchers may obtain BMAP cDNA clones from RESEARCH  
GENETICS. It should be noted that Bento Soares is generating a  
small number of additional specialized non-redundant arrays of BMAP  
cDNAs whose availability will be considered under appropriate and  
limited collaborative arrangements  
Seq primer: M13 Forward  
PDIAs=es.

## FEATURES

## source

```
Location/Qualifiers
1. .318
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UI-M-CG0p-bnr-h-05-0-UI"
/lab_host="DH10B (Life Technologies)"
/clone_lib="NIH_BMAP Ret4_S2"
/notes="Vector: p7T3D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; The
NIH_BMAP Ret4_S2 library is a subtracted library,
ultimately derived from mouse retina tissue libraries at
various stages of development. For a detailed description
of the library from which this clone was derived, please
visit our web site at brainest.eng.uiowa.edu. The tissue
for this library was contributed by Dr. Xin-Yuan Fu, Yale
University School of Medicine
TAG_TISSUE=adult-retina
TAG_LIB=NIH_BMAP Ret4_S2
TAG_SEQ=GTGAGCGCGCAC"
```

## ORIGIN

```
Query Match      4.2%; Score 20; DB 10; Length 318;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 369 AATTATGTTTCATTATTATT 388
      |||||
Db 251 AATTATGTTTCATTATTATT 270
```

## RESULT 15

```
AA410171      331 bp mRNA linear EST 26-AUG-1998
LOCUS      AA410171      331 bp mRNA linear EST 26-AUG-1998
DEFINITION      EST01724 Mouse 7.5 dpc embryo ectoplacental cone cDNA library Mus
ACCESSION      AA410171
VERSION      AA410171.1 GI:2066639
KEYWORDS      EST.
SOURCE      Mus musculus (house mouse)
```

```
REFERENCE      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 331)
Ko, M.S.H., Threalt, T.A., Horton, J.H., Wang, X., Cui, Y., Wang, X.,
Pryor, E., Paris, J., Wells-Smith, J., Fujiwara, H., Yotsumoto, S. and
Nakashima, H.
Systematic analyses of mouse genes expressed in embryo implantation
site
```

```
JOURNAL      Unpublished (1997)
```

```
COMMENT      Other ESTs: EST01725
```

```
Contact: Ko, MSH
```

Center for Molecular Medicine and Genetics

Wayne State University

5047 Gullen Mall, Detroit, MI 48202

Tel: 3135776708

Fax: 3135776200

Email: [msko@cmb.biosci.wayne.edu](mailto:msko@cmb.biosci.wayne.edu)

```

FEATURES             Seq primer: M13 Forward.
  source             Location/Qualifiers
    1..331
      /organism="Mus musculus"
      /mol_type="mRNA"
      /strain="C57BL/6J"
      /db_xref="AtCC (inhost):1363847"
      /db_xref="taxon:10090"
      /clone="C0013C12"
      /sex="unknown"
      /tissue_type="ectoplacental cone"
      /dev_stage="embryonic day 7.5 postconception"
      /lab_host="DH10B"
      /clone_lib="Mouse 7.5 dpc embryo ectoplacental cone cDNA
      library"
      /note="Organ: embryo; Vector: pSPORT1 (Life Technologies);
      Site_1: SalI; Site_2: NotI; Total RNAs were extracted from
      ectoplacental cone of 7.5-dpc embryos. The
      double-stranded cDNA was synthesized from total RNAs with
      an Oligo(dt) primer. The library was constructed by Minoru
      S. H. Ko."

```

ORIGIN

```

Query Match          4.2%; Score 20; DB 9; Length 331;
Best Local Similarity 100.0%; Pred.No. 44;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  369 AATTATGTTTCATTATTATT 388
     |||||
DB   281 AATTATGTTTCATTATTATT 300

```

Search completed: July 31, 2004, 10:50:08  
 Job time : 2005 secs



anCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2004, 09:32:50 ; Search time 336 Seconds  
(without alignments)  
6946.115 Million cell updates/sec

Title: US-10-020-540A-1

Perfect score: 476

Sequence: 1 aactatgctgacgaagat.....aactccctggttacaacc 476

Scoring table: OLIGO NUC

Gapop 60.0 , Gapext 60.0

Searched: 3222919 seqs, 245157024 residues

Word size : 0

Total number of hits satisfying chosen parameters: 6445838

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Published Applications NA:

- 1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*
- 2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq.\*
- 3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*
- 4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq.\*
- 5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*
- 6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq.\*
- 7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*
- 8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*
- 9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq.\*
- 10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq.\*
- 11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq.\*
- 12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*
- 13: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq.\*
- 14: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq.\*
- 15: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq.\*
- 16: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq.\*
- 17: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*
- 18: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*
- 19: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	476	100.0	476	US-10-020-540A-1	Sequence 1, Appli
2	476	100.0	518	US-10-020-540A-2	Sequence 2, Appli
3	369	77.5	651	US-10-020-540A-3	Sequence 3, Appli
4	369	77.5	701	US-10-020-540A-4	Sequence 4, Appli
5	21	4.4	373	US-10-424-599-104814	Sequence 104814,
6	21	4.4	622	US-10-027-632-229768	Sequence 229768,
7	21	4.4	622	US-10-027-632-229768	Sequence 229768,
8	20	4.2	9973	US-10-085-117-109	Sequence 109, App
9	19	4.0	1553	US-10-437-963-22143	Sequence 22143, A
10	19	4.0	3237	US-10-437-963-22139	Sequence 22139, A
11	19	4.0	5675	US-10-369-493-29305	Sequence 29305, A
12	18	3.8	28	US-10-020-540A-10	Sequence 10, Appl
13	18	3.8	327	US-09-867-701-603	Sequence 603, App
14	18	3.8	402	US-10-437-963-76846	Sequence 76846, A

15	18	3.8	454	9	US-09-867-701-949	Sequence 949, App
16	18	3.8	536	9	US-09-867-701-733	Sequence 733, App
c 17	18	3.8	553	13	US-10-424-599-1806	Sequence 1806, A
c 18	18	3.8	570	13	US-10-027-632-184726	Sequence 184726,
c 19	18	3.8	570	16	US-10-027-632-184726	Sequence 184726,
20	18	3.8	637	13	US-10-027-632-211784	Sequence 211784,
21	18	3.8	637	16	US-10-027-632-211784	Sequence 211784,
22	18	3.8	668	13	US-10-027-632-260421	Sequence 260421,
23	18	3.8	668	16	US-10-027-632-260421	Sequence 260421,
c 24	18	3.8	872	15	US-10-029-386-22974	Sequence 22974, A
c 25	18	3.8	1752	17	US-10-437-963-54963	Sequence 54963, A
c 26	18	3.8	2259	13	US-10-027-632-255809	Sequence 255809,
c 27	18	3.8	2259	16	US-10-027-632-255810	Sequence 255810,
c 28	18	3.8	2259	16	US-10-027-632-255809	Sequence 255809,
c 29	18	3.8	2259	16	US-10-027-632-255810	Sequence 255810,
c 30	18	3.8	2597	15	US-10-037-270-48	Sequence 48, Appl
c 31	18	3.8	2597	16	US-10-117-722-48	Sequence 48, Appl
c 32	18	3.8	7380	13	US-09-939-853A-10	Sequence 10, Appl
c 33	18	3.8	7380	16	US-09-939-853A-12	Sequence 12, Appl
c 34	18	3.8	16001	15	US-10-085-959-21	Sequence 21, Appl
c 35	18	3.8	38584	15	US-10-114-170-50	Sequence 50, Appl
c 36	18	3.8	63686	13	US-10-087-192-466	Sequence 466, App
c 37	18	3.8	79799	13	US-10-087-192-346	Sequence 346, App
c 38	18	3.8	138115	17	US-10-322-281-377	Sequence 377, App
c 39	17	3.6	162	13	US-10-085-783A-49002	Sequence 49002, A
c 40	17	3.6	182	16	US-10-242-535A-49002	Sequence 49002, A
c 41	17	3.6	231	10	US-09-535-459-2040	Sequence 2040, App
c 42	17	3.6	247	9	US-09-983-965-2222	Sequence 2222, App
c 43	17	3.6	270	10	US-09-535-459-2044	Sequence 2044, App
c 44	17	3.6	295	9	US-09-294-093B-233	Sequence 233, App
c 45	17	3.6	302	10	US-09-535-459-1934	Sequence 1934, App

ALIGNMENTS

RESULT 1

US-10-020-540A-1  
; Sequence 1, Application US/10020540A  
; Publication No. US20020182593A1  
; GENERAL INFORMATION:  
; APPLICANT: Wu, Gusui  
; APPLICANT: McBrady, Kevin  
; TITLE OF INVENTION: Strawberry Vein Banding Virus Promoter  
; TITLE OF INVENTION: (SVEB) Promoter  
; FILE REFERENCE: 0173.210us  
; CURRENT APPLICATION NUMBER: US/10/020,540A  
; CURRENT FILING DATE: 2001-10-30  
; PRIOR APPLICATION NUMBER: 60/245,354  
; PRIOR FILING DATE: 2000-11-01  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 476  
; TYPE: DNA  
; ORGANISM: Strawberry Vein Banding Virus (SVEB) Strain E3  
US-10-020-540A-1

Query Match	100.0%	Score 476;	DB 14;	Length 476;
Best Local Similarity	100.0%	Pred. No. 7.1e-240;		
Matches 476;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	AACATATGCTGATGACAGATAATCTTAATAGCAATATTTCAGAAATTATCAAGAGAGAAA	60	
Db	1	AACATATGCTGATGACAGATAATCTTAATAGCAATATTTCAGAAATTATCAAGAGAGAAA	60	
QY	61	GAATTAATACTCTTTTTCAGAAATATGAGCCGCTTTTACAGTGCCGAGCTAGTATCACT	120	
Db	61	GAATTAATACTCTTTTTCAGAAATATGAGCCGCTTTTACAGTGCCGAGCTAGTATCACT	120	
QY	121	GAAGAAGACAGACAGACAAATGCTGCTCGATGACACAGAACCAATCTTTGCACGATGT	180	
Db	121	GAAGAAGACAGACAGACAAATGCTGCTCGATGACACAGAACCAATCTTTGCACGATGT	180	

QY	181	GAAGCAGCCAGAGTGGTCCCAAGACGCATCTCAGAAAAGGCATCTTCTTACCGACACAGAA	240
Db	181	GAAGCAGCCAGAGTGGTCCCAAGACGCATCTCAGAAAAGGCATCTTCTTACCGACACAGAA	240
QY	241	AAAGCAACCAAGCTCATCATCCAAATGATGACTGTCTGTTATGCGTCGGCTGAAGATA	300
Db	241	AAAGCAACCAAGCTCATCATCCAAATGATGACTGTCTGTTATGCGTCGGCTGAAGATA	300
QY	301	AGACTGACCCAGCCAGCCAGCACTAAAGAAAGAAATAATGCAAGTGGTCCTAGCTCCACTTTA	360
Db	301	AGACTGACCCAGCCAGCCAGCACTAAAGAAAGAAATAATGCAAGTGGTCCTAGCTCCACTTTA	360
QY	361	GCTTTAAATAATATGTTTTCACTATTATTTCTCTGCTTTTGTCTCTCTATATAAAGAGCTGTG	420
Db	361	GCTTTAAATAATATGTTTTCACTATTATTTCTCTGCTTTTGTCTCTCTATATAAAGAGCTGTG	420
QY	421	ATTTTCTATTTGAAGGCAGAGCGCAACACACACAGAAACCTCCCTGCTTACAAACC	476
Db	421	ATTTTCTATTTGAAGGCAGAGCGCAACACACACAGAAACCTCCCTGCTTACAAACC	476

## RESULT 2

```

US-10-020-540A-2
; Sequence 2, Application US/10020540A
; Publication NO. US29020182593A1
; GENERAL INFORMATION:
; APPLICANT: WU, Gusui
; APPLICANT: McBryde, Kevin
; TITLE OF INVENTION: Strawberry Vein Banding Virus Promoter
; TITLE OF INVENTION: (SBBV) Promoter
; FILE REFERENCE: 0173.210us
; CURRENT APPLICATION NUMBER: US/10/020,540A
; CURRENT FILING DATE: 2001-10-30
; PRIOR APPLICATION NUMBER: 60/245,354
; PRIOR FILING DATE: 2000-11-01
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 518
; TYPE: DNA
; ORGANISM: Strawberry Vein Banding Virus (SBBV) Strain E3
US-10-020-540A-2

```

Db	375	GCTTTAATAATTAAGTTTCATTATTATTCTCTGCTTTTGCCTCTATATAAAGAGGCTGT	434
Qy	421	ATTTTCATTTGAAGGAGAGGGCAACACACACAGAACTCCCTGCTTTACAACC	476
Db	435	ATTTTCATTTGAAGGAGAGGGCAACACACACAGAACTCCCTGCTTTACAACC	490

### RESULT 3

```

US-10-020-540A-3
; Sequence 3, Application US/10020540A
; Publication No. US20020182593A1
; GENERAL INFORMATION:
;
; APPLICANT: Wu, Gusui
; APPLICANT: Kevin
; APPLICANT: McBryde, Kevin
; TITLE OF INVENTION: Strawberry Vein Banding Virus Promoter
; TITLE OF INVENTION: (SVBV) Promoter
; FILE REFERENCE: 0173.210us
; CURRENT APPLICATION NUMBER: US/10/020,540A
; CURRENT FILING DATE: 2001-10-30
; PRIOR APPLICATION NUMBER: 60/245,354
; PRIOR FILING DATE: 2000-11-01
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 3
;
; LENGTH: 651
; TYPE: DNA
; ORGANISM: Strawberry Vein Banding Virus (SVBV) Strain E3
US-10-020-540A-3

```

## RESULT 4

```

RESUME 4
US-10-020-540A-4
; Sequence 4, Application US/10020540A
; Publication No. US20020182593A1
; GENERAL INFORMATION:
; APPLICANT: Wu, Gusui
; APPLICANT: McBryde, Kevin
; TITLE OF INVENTION: Strawberry Vein Banding Virus Promoter
; TITLE OF INVENTION: (SVEV) Promoter
; FILE REFERENCE: 0173.210us

```

CURRENT APPLICATION NUMBER: US/10/020,540A  
CURRENT FILING DATE: 2001-10-30  
PRIOR APPLICATION NUMBER: 60/245,354  
PRIOR FILING DATE: 2000-11-01  
NUMBER OF SEQ ID NOS: 10  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 4  
LENGTH: 701  
TYPE: DNA  
ORGANISM: Strawberry Vein Banding Virus (SVBV) Strain E3  
US-10-020-540A-4

Query Match 77.5%; Score 369; DB 14; Length 701;  
Best Local Similarity 100.0%; Pred. No. 1.6e-183; Indels 0; Gaps 0;  
Matches 369; Conservative 0; Mismatches 0; Gaps 0;

QY 108 GCTAGCTATCACTGAAAGACAGACAGACAATGGTGTCTCGATGACACAGAACCACTCT 167  
DB 319 GCTAGCTATCACTGAAAGACAGACAGACAATGGTGTCTCGATGACACAGAACCACTCT 378  
QY 168 TTGCACGATGTGAAGCGACGAGTGTCCACAGACGCACTCAGAAAAGGCATCTTC 227  
DB 379 TTGCACGATGTGAAGCGACGAGTGTCCACAGACGCACTCAGAAAAGGCATCTTC 438  
QY 228 TACCGACAGAAAAAGACAAACACAGCTCATCATCAATGTAGACTGTGTTATGG 287  
DB 439 TACCGACAGAAAAAGACAAACACAGCTCATCATCAATGTAGACTGTGTTATGG 498  
QY 288 TCGGCTGAAGATAAGACTGACCCAGCCAGCACTAAAGAGAAATATGCAAGTGGTCC 347  
DB 499 TCGGCTGAAGATAAGACTGACCCAGCCAGCACTAAAGAGAAATATGCAAGTGGTCC 558  
QY 348 TAGCTCCACTTTAGCTTTAATATTATTTGTTTCAATTTATTTCTCTCTCTCTAT 407  
DB 559 TAGCTCCACTTTAGCTTTAATATTATTTGTTTCAATTTATTTCTCTCTCTCTAT 618  
QY 408 ATAAAGAGCTTGTATTTTTCATTTTGAAGGCGAGCGGACACACACAGAACTCCCTGC 467  
DB 619 ATAAAGAGCTTGTATTTTTCATTTTGAAGGCGAGCGGACACACACAGAACTCCCTGC 678  
QY 468 TTACAAACC 476  
DB 679 TTACAAACC 687

RESULT 5  
US-10-424-599-104814  
Sequence 104814, Application US/10424599  
Publication No. US20040031072A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa Thomas J  
APPLICANT: Kovalic David K  
APPLICANT: Zhou Yihua  
APPLICANT: Cao Yongwei  
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
FILE REFERENCE: 38-21(53223)B  
CURRENT APPLICATION NUMBER: US/10/424,599  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 285684  
SEQ ID NO 104814  
LENGTH: 373  
TYPE: DNA  
ORGANISM: Glycine max  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT3847\_65663C.1  
US-10-424-599-104814

Query Match 4.4%; Score 21; DB 13; Length 373;  
Best Local Similarity 100.0%; Pred. No. 2.4; Indels 0; Gaps 0;  
Matches 367; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 367 ATAATTATGTTTCATTATTAT 387

DB 213 ATAATTATGTTTCATTATTAT 233  
RESULT 6  
US-10-027-632-229768/c  
Sequence 229768, Application US/10027632  
Publication No. US20020198371A1  
GENERAL INFORMATION:  
APPLICANT: Wang, David G.  
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
FILE REFERENCE: 108827,129  
CURRENT APPLICATION NUMBER: US/10/027,632  
CURRENT FILING DATE: 2002-04-30  
PRIOR APPLICATION NUMBER: US 60/218,006  
PRIOR FILING DATE: 2000-07-12  
PRIOR APPLICATION NUMBER: US 60/198,676  
PRIOR FILING DATE: 2000-04-20  
PRIOR APPLICATION NUMBER: US 60/193,483  
PRIOR FILING DATE: 2000-03-29  
PRIOR APPLICATION NUMBER: US 60/185,218  
PRIOR FILING DATE: 2000-02-24  
PRIOR APPLICATION NUMBER: US 60/167,363  
PRIOR FILING DATE: 1999-11-23  
PRIOR APPLICATION NUMBER: US 60/156,358  
PRIOR FILING DATE: 1999-09-28  
PRIOR APPLICATION NUMBER: US 60/146,002  
PRIOR FILING DATE: 1999-08-09  
NUMBER OF SEQ ID NOS: 325720  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 229768  
LENGTH: 622  
TYPE: DNA  
ORGANISM: Human  
US-10-027-632-229768

Query Match 4.4%; Score 21; DB 13; Length 622;  
Best Local Similarity 100.0%; Pred. No. 2.5; Indels 0; Gaps 0;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 322 TAAAGAGAAATATGCAAGT 342  
DB 545 TAAAGAGAAATATGCAAGT 525

RESULT 7  
US-10-027-632-229768/c  
Sequence 229768, Application US/10027632  
Publication No. US20030204075A9  
GENERAL INFORMATION:  
APPLICANT: Wang, David G.  
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
FILE REFERENCE: 108827,129  
CURRENT APPLICATION NUMBER: US/10/027,632  
CURRENT FILING DATE: 2002-04-30  
PRIOR APPLICATION NUMBER: US 60/218,006  
PRIOR FILING DATE: 2000-07-12  
PRIOR APPLICATION NUMBER: US 60/198,676  
PRIOR FILING DATE: 2000-04-20  
PRIOR APPLICATION NUMBER: US 60/193,483  
PRIOR FILING DATE: 2000-03-29  
PRIOR APPLICATION NUMBER: US 60/185,218  
PRIOR FILING DATE: 2000-02-24  
PRIOR APPLICATION NUMBER: US 60/167,363  
PRIOR FILING DATE: 1999-11-23  
PRIOR APPLICATION NUMBER: US 60/156,358  
PRIOR FILING DATE: 1999-09-28  
PRIOR APPLICATION NUMBER: US 60/146,002  
PRIOR FILING DATE: 1999-08-09  
NUMBER OF SEQ ID NOS: 325720  
SOFTWARE: FastSeq for Windows Version 4.0

```
; SEQ ID NO 229768
; LENGTH: 622
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-229768

Query Match
Best Local Similarity 4.4%; Score 21; DB 16; Length 622;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 322 TAAGAAGAAATAATGCAAGT 342
Db 545 TAAGAAGAAATAATGCAAGT 525

RESULT 8
US-10-085-117-109/c
; Sequence 109, Application US/10085117
; Publication No. US20030232334A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER
; FILE REFERENCE: 529452000121
; CURRENT APPLICATION NUMBER: US/10/085,117
; CURRENT FILING DATE: 2002-02-27
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 361
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 109
; LENGTH: 99973
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: Variation
; LOCATION: (1)...(99973)
; OTHER INFORMATION: n = any nucleotide
US-10-085-117-109

Query Match
Best Local Similarity 4.2%; Score 20; DB 16; Length 99973;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 421 ATTTCATTGAAGGCAGAG 440
Db 76871 ATTTCATTGAAGGCAGAG 76852

RESULT 9
US-10-437-963-22143/c
; Sequence 22143, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: RICE NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
; TITLE OF INVENTION: PLANTS AND USES THEREOF FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 22143
; LENGTH: 1553
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
```

```
; OTHER INFORMATION: Clone ID: PAT_MRT4530_27347C.1
US-10-437-963-22143

Query Match
Best Local Similarity 4.0%; Score 19; DB 17; Length 1553;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 155 CAGAACCAATCTTTGCAG 173
Db 1110 CAGAACCAATCTTTGCAG 1092

RESULT 10
US-10-437-963-22139/c
; Sequence 22139, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: RICE NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH
; TITLE OF INVENTION: PLANTS AND USES THEREOF FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 22139
; LENGTH: 3237
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_27343C.1
US-10-437-963-22139

Query Match
Best Local Similarity 4.0%; Score 19; DB 17; Length 3237;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 155 CAGAACCAATCTTTGCAG 173
Db 2536 CAGAACCAATCTTTGCAG 2518

RESULT 11
US-10-369-493-29305/c
; Sequence 29305, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 29305
; LENGTH: 5675
; TYPE: DNA
; ORGANISM: Caenorhabditis elegans
US-10-369-493-29305

Query Match
Best Local Similarity 4.0%; Score 19; DB 16; Length 5675;
```

Best Local Similarity 100.0%; Pred. No. 35;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 437 AGAGGCGAACACACACACA 455  
Db 3868 AGAGGCGAACACACACACA 3850

## RESULT 12

US-10-020-540A-10/c  
; Sequence 10, Application US/10020540A  
; Publication No. US20020182593A1  
; GENERAL INFORMATION:  
; APPLICANT: Wu, Gusui  
; APPLICANT: McBryde, Kevin  
; TITLE OF INVENTION: Strawberry Vein Banding Virus Promoter  
; TITLE OF INVENTION: (SVBV) Promoter  
; FILE REFERENCE: 0173.210us  
; CURRENT APPLICATION NUMBER: US/10/020,540A  
; CURRENT FILING DATE: 2001-10-30  
; PRIOR APPLICATION NUMBER: 60/245,354  
; PRIOR FILING DATE: 2000-11-01  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 10  
; LENGTH: 28  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Primer  
US-10-020-540A-10

Query Match 3.8%; Score 18; DB 14; Length 28;  
Best Local Similarity 100.0%; Pred. No. 70;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 459 CCTCCCTGCTTACAAACC 476  
Db 28 CCTCCCTGCTTACAAACC 11

## RESULT 13

US-09-867-701-603/c  
; Sequence 603, Application US/09867701  
; Patent No. US20020132237A1  
; GENERAL INFORMATION:  
; APPLICANT: Aglate, Paul A.  
; APPLICANT: Jones, Robert  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER  
; FILE REFERENCE: 210121.497  
; CURRENT APPLICATION NUMBER: US/09/867,701  
; CURRENT FILING DATE: 2001-05-29  
; NUMBER OF SEQ ID NOS: 10912  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 603  
; LENGTH: 327  
; TYPE: DNA  
; ORGANISM: Homo sapien  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(327)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-867-701-603

Query Match 3.8%; Score 18; DB 9; Length 327;  
Best Local Similarity 100.0%; Pred. No. 89;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 77 CAGAATATGAAGCCCGCT 94  
Db 153 CAGAATATGAAGCCCGCT 136

## RESULT 14

US-10-437-963-76846/c  
; Sequence 76846, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 76846  
; LENGTH: 402  
; TYPE: DNA  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_76802C.1  
US-10-437-963-76846

Query Match 3.8%; Score 18; DB 17; Length 402;  
Best Local Similarity 100.0%; Pred. No. 91;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 179 GTGAAGCAGCCAGAGTGG 196  
Db 259 GTGAAGCAGCCAGAGTGG 242

## RESULT 15

US-09-867-701-949  
; Sequence 949, Application US/09867701  
; Patent No. US20020132237A1  
; GENERAL INFORMATION:  
; APPLICANT: Aglate, Paul A.  
; APPLICANT: Jones, Robert  
; APPLICANT: Harlocker, Susan L.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER  
; FILE REFERENCE: 210121.497  
; CURRENT APPLICATION NUMBER: US/09/867,701  
; CURRENT FILING DATE: 2001-05-29  
; NUMBER OF SEQ ID NOS: 10912  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 949  
; LENGTH: 454  
; TYPE: DNA  
; ORGANISM: Homo sapien  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(454)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-867-701-949

Query Match 3.8%; Score 18; DB 9; Length 454;  
Best Local Similarity 100.0%; Pred. No. 92;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 77 CAGAATATGAAGCCCGCT 94  
Db 140 CAGAATATGAAGCCCGCT 157

Search completed: July 31, 2004, 10:57:04

Mon Aug 2 10:06:59 2004

us-10-020-540a-1.olig.rnpb

Page 6

Job time : 338 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2004, 09:31:35 ; Search time 65 Seconds  
(without alignments)  
4063.948 Million cell updates/sec

Title: US-10-020-540A-1

Perfect score: 476

Sequence: 1 aactatgctgatgacaagat.....aacctccctgcttacaacc 476

Scoring table: OLIGO\_NUC

Gapop\_60.0 , Gapext 60.0

Searched: 682709 seqs, 277475446 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued Patents NA:\*

- 1: /cgn2\_6/ptodata/2/ina/5A\_COMB.seq:\*
- 2: /cgn2\_6/ptodata/2/ina/5B\_COMB.seq:\*
- 3: /cgn2\_6/ptodata/2/ina/6A\_COMB.seq:\*
- 4: /cgn2\_6/ptodata/2/ina/6B\_COMB.seq:\*
- 5: /cgn2\_6/ptodata/2/ina/PTUS\_COMB.seq:\*
- 6: /cgn2\_6/ptodata/2/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	DB ID	Description
C 1	18	3.8	2597 4	US-09-620-312D-48
C 2	18	3.8	38584 4	US-09-453-702B-50
C 3	17	3.6	501 4	US-09-543-681A-597
C 4	17	3.6	528 4	US-09-401-064-275
C 5	17	3.6	726 4	US-09-107-532A-1355
C 6	17	3.6	1566 4	US-09-489-039A-947
C 7	17	3.6	1975 4	US-09-308-345A-6
C 8	17	3.6	3562 3	US-09-360-197-1
C 9	17	3.6	4105 4	US-09-589-567-1
C 10	17	3.6	4617 4	US-09-930-055A-1
C 11	17	3.6	5109 4	US-09-930-055A-2
C 12	16	3.4	298 4	US-09-313-294A-4713
C 13	16	3.4	440 4	US-09-621-976-2230
C 14	16	3.4	813 4	US-09-134-001C-2724
C 15	16	3.4	825 4	US-09-956-171E-1111
C 16	16	3.4	966 4	US-09-134-001C-53
C 17	16	3.4	1113 3	US-09-172-353-1
C 18	16	3.4	1113 4	US-09-793-955-1
C 19	16	3.4	1116 3	US-09-776-971-139
C 20	16	3.4	1221 4	US-09-134-001C-365
C 21	16	3.4	1591 3	US-09-468-856B-9
C 22	16	3.4	1591 3	US-09-468-856B-9
C 23	16	3.4	2013 4	US-09-134-000C-1201
C 24	16	3.4	2813 4	US-09-689-255C-3
C 25	16	3.4	3043 3	US-09-008-271A-14
C 26	16	3.4	3332 4	US-09-448-806C-1
C 27	16	3.4	3401 4	US-09-553-554-3

C 28	16	3.4	3441 4	US-09-095-881-1
C 29	16	3.4	3502 2	US-08-724-394A-16
C 30	16	3.4	4080 4	US-09-016-434-1292
C 31	16	3.4	4848 4	US-08-961-527-185
C 32	16	3.4	10952 1	US-08-602-036A-1
C 33	16	3.4	10952 2	US-08-502-374A-1
C 34	16	3.4	10952 2	US-08-642-407A-1
C 35	16	3.4	24979 2	US-08-147-777-3
C 36	16	3.4	24979 3	US-08-452-872-3
C 37	16	3.4	24979 5	PCT-US93-03985-3
C 38	16	3.4	152331 3	US-09-128-155-16
C 39	16	3.4	176373 3	US-09-128-155-17
C 40	16	3.4	640681 4	US-09-730-988-1
C 41	16	3.4	640681 4	US-09-750-988-1
C 42	16	3.4	1230025 4	US-09-198-452A-1
C 43	16	3.4	1664976 4	US-08-916-421B-1
C 44	15	3.2	88 6	5198347-12
C 45	15	3.2	189 4	US-09-134-000C-2505

ALIGNMENTS

RESULT 1

US-09-620-312D-48/c  
; Sequence 48, Application US/09620312D  
; Patent No. 6569662  
; GENERAL INFORMATION:  
; APPLICANT: Tang, Y. Tom  
; APPLICANT: Liu, Chenghua  
; APPLICANT: Asundi, Vinod  
; APPLICANT: Zhang, Jie  
; APPLICANT: Ren, Feiyen  
; APPLICANT: Chen, Rui-hong  
; APPLICANT: Zhao, Qing A.  
; APPLICANT: Wehrman, Tom  
; APPLICANT: Xue, Aidong J.  
; APPLICANT: Wang, Yonghong  
; APPLICANT: Wang, Jian-Rui  
; APPLICANT: Zhou, Ping  
; APPLICANT: Ma, Yunqing  
; APPLICANT: Wang, Dunrui  
; APPLICANT: Wang, Zhiwei  
; APPLICANT: John Tillinghast  
; APPLICANT: Drmanac, Radoje T.  
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and  
; TITLE OF INVENTION: Polypeptides  
; FILE REFERENCE: 784CIP2B  
; CURRENT APPLICATION NUMBER: US/09/620,312D  
; CURRENT FILING DATE: 2000-07-19  
; PRIOR APPLICATION NUMBER: 09/552,317  
; PRIOR FILING DATE: 2000-04-25  
; PRIOR APPLICATION NUMBER: 09/488,725  
; PRIOR FILING DATE: 2000-01-21  
; NUMBER OF SEQ ID NOS: 1105  
; SOFTWARE: pt\_FL\_Genes Version 1.0  
; SEQ ID NO 48  
; LENGTH: 2597  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (592)..(1272)  
US-09-620-312D-48

Query Match 3.8%; Score 18; DB 4; Length 2597;

Best Local Similarity 100.0%; Pred. No. 12;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 77 CAGAAATATGAGCCCGCT 94

Db 1658 CAGAAATATGAGCCCGCT 1641

```
RESULT 2
US-09-453-702B-50
; Sequence 50, Application US/09453702B
; Patent No. 6365723
; GENERAL INFORMATION:
; APPLICANT: Blattner, Frederick R.
; Burlaud, Valerie
; Perna, Nicole T.
; Plunkett, Guy
; Welch, Rod
; TITLE OF INVENTION: No. 6365723el Sequences of E. coli O157
; NUMBER OF SEQUENCES: 265
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Charles & Brady
; STREET: 1 South Pinckney Street
; CITY: Madison
; STATE: WI
; COUNTRY: US
; ZIP: 53701-2113
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.44mb storage
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 8.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/453,702B
; FILING DATE: 03-Dec-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/110,955
; FILING DATE: 04-Dec-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Seay, Nicholas J.
; REGISTRATION NUMBER: 27386
; REFERENCE/DOCKET NUMBER: 960296.95017
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (608) 251-5000
; TELEFAX: (608) 251-9166
; INFORMATION FOR SEQ ID NO: 50:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 38584
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; SEQUENCE DESCRIPTION: SEQ ID NO: 50:
US-09-453-702B-50
Query Match 3.8%; Score 18; DB 4; Length 38584;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 77 CAGAAATGAGCCCGCT 94
Db 3180 CAGAAATGAGCCCGCT 3197
|||||

RESULT 3
US-09-543-681A-597/c
; Sequence 597, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543.681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 597
```

```
; LENGTH: 501
; TYPE: DNA
; ORGANISM: Proteus mirabilis
US-09-543-681A-597
Query Match 3.6%; Score 17; DB 4; Length 501;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 408 ATAAAGAGCTTGATTT 424
Db 106 ATAAAGAGCTTGATTT 90
|||||

RESULT 4
US-09-401-064-275/c
; Sequence 275, Application US/09401064
; Patent No. 6623923
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Lodes, Michael J.
; APPLICANT: Secrist, Heather
; APPLICANT: Benson, Darin R.
; APPLICANT: Meagher, Madeline Joy
; APPLICANT: Stolk, John A.
; APPLICANT: Wang, Tongtong
; TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF COLON CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.471C2
; CURRENT APPLICATION NUMBER: US/09/401,064
; CURRENT FILING DATE: 1999-09-22
; NUMBER OF SEQ ID NOS: 371
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 275
; LENGTH: 528
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)..(528)
; OTHER INFORMATION: n = A,T,C or G
US-09-401-064-275
Query Match 3.6%; Score 17; DB 4; Length 528;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 123 AAAGACAGCAAGCAAT 139
Db 391 AAAGACAGCAAGCAAT 375
|||||

RESULT 5
US-09-107-532A-1355
; Sequence 1355, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
```



```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Arinello, Pamela Deneké
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 1355:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 726 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...726
; SEQUENCE DESCRIPTION: SEQ ID NO: 1355:
US-09-107-532A-1355

```

```

Query Match 3.6%; Score 17; DB 4; Length 726;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 325 AGAAGAAATAATCAAG 341
Db 18 AGAAGAAATAATCAAG 34

```

```

RESULT 6
US-09-489-039A-947
; Sequence 947, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR FILING DATE: 1999-01-29
; PRIOR APPLICATION NUMBER: US 60/117,747
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 947
; LENGTH: 1566
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-947

```

```

Query Match 3.6%; Score 17; DB 4; Length 1566;
Best Local Similarity 100.0%; Pred. No. 39;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 337 GCAAGTGGCTCTAGTCT 353
Db 348 GCAAGTGGCTCTAGTCT 364

```

```

RESULT 7
US-09-308-345A-6
; Sequence 6, Application US/09308345A

```

```

; Patent No. 6569665
; GENERAL INFORMATION:
; APPLICANT: BOEHM, Thomas;
; APPLICANT: DEAR, Neil T.
; TITLE OF INVENTION: No. 6569665el calpains, their preparation and use
; FILE REFERENCE: 0050/47576
; CURRENT APPLICATION NUMBER: US/09/308,345A
; CURRENT FILING DATE: 1999-05-19
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: WordPerfect v. 6.1
; SEQ ID NO 6
; LENGTH: 1975
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 44...1963
US-09-308-345A-6

```

```

Query Match 3.6%; Score 17; DB 4; Length 1975;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 216 AAAGGCATCTTCTACCG 232
Db 1736 AAAGGCATCTTCTACCG 1752

```

```

RESULT 8
US-09-360-197-1
; Sequence 1, Application US/09360197
; Patent No. 6287859
; GENERAL INFORMATION:
; APPLICANT: Basilana, Frederic
; APPLICANT: Lazdunski, Michel
; APPLICANT: Waldmann, Rainer
; APPLICANT: Deweille, Jan R.
; TITLE OF INVENTION: Human and Rat Families of Neuronal Acid-Sensitive
; FILE REFERENCE: 989.6708P
; CURRENT APPLICATION NUMBER: US/09/360,197-
; CURRENT FILING DATE: 1997-07-23
; PRIOR APPLICATION NUMBER: 09/129,758
; PRIOR FILING DATE: 1998-08-05
; PRIOR APPLICATION NUMBER: 60/095,408
; PRIOR FILING DATE: 1998-08-05
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 3562
; TYPE: DNA
; ORGANISM: rattus sp.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (123)..(1700)
US-09-360-197-1

```

```

Query Match 3.6%; Score 17; DB 3; Length 3562;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 445 ACACACACACAGAACT 461
Db 36 ACACACACACAGAACT 52

```

```

RESULT 9
US-09-589-567-1/c
; Sequence 1, Application US/09589567
; Patent No. 6479730
; GENERAL INFORMATION:
; APPLICANT: Mahajan, Pramod B.
; APPLICANT: Shi, Jinrui

```

;; TITLE OF INVENTION: Maize DNA Ligase II Orthologue and Uses  
;; FILE OF INVENTION: Therecf  
;; FILE REFERENCE: 1125  
;; CURRENT APPLICATION NUMBER: US/09/589,567  
;; PRIOR FILING DATE: 2000-06-07  
;; PRIOR APPLICATION NUMBER: US 60/145,911  
;; PRIOR FILING DATE: 1999-07-27  
;; NUMBER OF SEQ ID NOS: 3  
;; SOFTWARE: FastSeq for Windows Version 3.0  
;; SEQ ID NO 1  
;; LENGTH: 4105  
;; TYPE: DNA  
;; ORGANISM: Zea Mays  
;; FEATURE:  
;; NAME/KEY: CDS  
;; LOCATION: (148)...(3801)  
;; NAME/KEY: misc feature  
;; LOCATION: (2713)...(2745)  
;; OTHER INFORMATION: ATP-dependent DNA ligase signature sequence  
US-09-589-567-1

Query Match 3.6%; Score 17; DB 4; Length 4105;  
Best Local Similarity 100.0%; Pred. No. 41;  
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 158 AACCATCTTTGACG 174

Db 915 AACCATCTTTGACG 899

## RESULT 10

US-08-930-055A-1

;; Sequence 1, Application US/08930055A

;; Patent No. 6403561

;; GENERAL INFORMATION:

;; APPLICANT: ROSE, CHRISTINANE; VARGAS,

;; APPLICANT: FROYLAN; BOURGAT, PIERRE; SCHWARTZ, JEAN-

;; APPLICANT: CHARLES; BISHOP, PAUL; BAMBAL, RAMESH;

;; APPLICANT: GANELLIN, CHARON; LEBLOND, BERTRAND; MOORE,

;; APPLICANT: ANDREW; CHAN, SUZANNE; ZHAO, LIHUA.

;; TITLE OF INVENTION: TRIPEPTIDYLPEPTIDASE

;; NUMBER OF SEQUENCES: 8

;; CORRESPONDENCE ADDRESS:

;; ADDRESSEE: BIERMAN, MUSERLIAN AND LUCAS

;; ADDRESSEE: LLP

;; STREET: 600 THIRD AVENUE

;; CITY: NEW YORK

;; STATE: NEW YORK

;; COUNTRY: USA

;; ZIP: 10016

;; COMPUTER READABLE FORM:

;; MEDIUM TYPE: FLOPPY DISK

;; COMPUTER: IBM PC COMPATIBLE

;; OPERATING SYSTEM: PC-DOS/MSDOS

;; SOFTWARE: MICROSOFT WORD 97

;; CURRENT APPLICATION DATA:

;; APPLICATION NUMBER: US/08/930,055A

;; FILING DATE: 19-FEB-1998

;; PRIOR APPLICATION DATA:

;; APPLICATION NUMBER: PCT/FR96/00700

;; FILING DATE: 9-MAY-1996

;; PRIOR APPLICATION DATA:

;; APPLICATION NUMBER: FR95/05489

;; FILING DATE: 9-MAY-1995

;; ATTORNEY/AGENT INFORMATION:

;; NAME: CHARLES A. MUSERLIAN

;; REGISTRATION NUMBER: 19,683

;; REFERENCE/DOCKET NUMBER: 408.013

;; TELECOMMUNICATION INFORMATION:

;; TELEPHONE: E (212) 661-8000

;; TELEFAX: (212) 661-8002

;; TELEX:

;; INFORMATION FOR SEQ ID NO: 1:  
;; SEQUENCE CHARACTERISTICS:

;; LENGTH: 4617

;; TYPE: NUCLEIC ACID

;; STRANDEDNESS: DOUBLE

;; TOPOLOGY: LINEAR

;; MOLECULE TYPE: cDNA

;; ORIGINAL SOURCE:

;; ORGANISM: RAT

;; STRAIN:

;; INDIVIDUAL ISOLATE:

;; DEVELOPMENTAL STAGE:

;; HAPLOTYPE:

;; TISSUE TYPE:

;; CELL TYPE:

;; CELL LINE:

;; ORGANELLE:

;; US-08-930-055A-1

Query Match 3.6%; Score 17; DB 4; Length 4617;

Best Local Similarity 100.0%; Pred. No. 41;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 300 AAGACTGACCCAGGCC 316

Db 3359 AAGACTGACCCAGGCC 3375

## RESULT 11

US-08-930-055A-2

;; Sequence 2, Application US/08930055A

;; Patent No. 6403561

;; GENERAL INFORMATION:

;; APPLICANT: ROSE, CHRISTINANE; VARGAS,

;; APPLICANT: FROYLAN; BOURGAT, PIERRE; SCHWARTZ, JEAN-

;; APPLICANT: CHARLES; BISHOP, PAUL; BAMBAL, RAMESH;

;; APPLICANT: GANELLIN, CHARON; LEBLOND, BERTRAND; MOORE,

;; APPLICANT: ANDREW; CHAN, SUZANNE; ZHAO, LIHUA.

;; TITLE OF INVENTION: TRIPEPTIDYLPEPTIDASE

;; NUMBER OF SEQUENCES: 8

;; CORRESPONDENCE ADDRESS:

;; ADDRESSEE: BIERMAN, MUSERLIAN AND LUCAS

;; ADDRESSEE: LLP

;; STREET: 600 THIRD AVENUE

;; CITY: NEW YORK

;; STATE: NEW YORK

;; COUNTRY: USA

;; ZIP: 10016

;; COMPUTER READABLE FORM:

;; MEDIUM TYPE: FLOPPY DISK

;; COMPUTER: IBM PC COMPATIBLE

;; OPERATING SYSTEM: PC-DOS/MSDOS

;; SOFTWARE: MICROSOFT WORD 97

;; CURRENT APPLICATION DATA:

;; APPLICATION NUMBER: US/08/930,055A

;; FILING DATE: 19-FEB-1998

;; PRIOR APPLICATION DATA:

;; APPLICATION NUMBER: PCT/FR96/00700

;; FILING DATE: 9-MAY-1996

;; PRIOR APPLICATION DATA:

;; APPLICATION NUMBER: FR95/05489

;; FILING DATE: 9-MAY-1995

;; ATTORNEY/AGENT INFORMATION:

;; NAME: CHARLES A. MUSERLIAN

;; REGISTRATION NUMBER: 19,683

;; REFERENCE/DOCKET NUMBER: 408.013

;; TELECOMMUNICATION INFORMATION:

;; TELEPHONE: E (212) 661-8000

;; TELEFAX: (212) 661-8002

;; TELEX:

;; INFORMATION FOR SEQ ID NO: 2:

;; SEQUENCE CHARACTERISTICS:

```
/ LENGTH: 5109
/ TYPE: NUCLEIC ACID
/ STRANDEDNESS: DOUBLE
/ TOPOLOGY: LINEAR
/ MOLECULE TYPE: CDNA
/ ORIGINAL SOURCE:
/ ORGANISM: RAT
/ STRAIN:
/ INDIVIDUAL ISOLATE:
/ DEVELOPMENTAL STAGE:
/ HAPLOTYPE:
/ TISSUE TYPE:
/ CELL TYPE:
/ CELL LINE:
/ ORGANELLE:
/ US-08-930-055A-2

Query Match          3.6%; Score 17; DB 4; Length 5109;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 300 AGACTGACCCGAGGCC 316
Db 3851 AAGACTGACCCGAGGCC 3867

RESULT 12
US-09-313-294A-4713/c
/ Sequence 4713, Application US/09313294A
/ Patent No. 6476212
/ GENERAL INFORMATION:
/ APPLICANT: Lalgudi, Raghunath V.
/ APPLICANT: Ito, Laura Y.
/ APPLICANT: Sherman, Bradley K.
/ TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
/ FILE REFERENCE: PL-0017 US
/ CURRENT APPLICATION NUMBER: US/09/313,294A
/ CURRENT FILING DATE: 1999-05-14
/ NUMBER OF SEQ ID NOS: 7600
/ SOFTWARE: PERL Program
/ SEQ ID NO 4713
/ LENGTH: 298
/ TYPE: DNA
/ ORGANISM: Zea mays
/ FEATURE:
/ NAME/KEY: misc feature
/ OTHER INFORMATION: Incyte ID No. 6476212 700348913H1
/ NAME/KEY: unsure
/ LOCATION: 259, 269, 288
/ OTHER INFORMATION: a, t, c, g, or other
/ US-09-313-294A-4713

Query Match          3.4%; Score 16; DB 4; Length 298;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 165 TCTTTGCAGCAGATGT 180
Db 221 TCTTTGCAGCAGATGT 206

RESULT 13
US-09-621-976-2230
/ Sequence 2230, Application US/09621976
/ Patent No. 6639063
/ GENERAL INFORMATION:
/ APPLICANT: Dumas Milne Edwards, J.B.
/ APPLICANT: Jobert, S.
/ APPLICANT: Giordano, J.Y.
/ TITLE OF INVENTION: ESTs and Encoded Human Proteins.
/ FILE REFERENCE: GENSET.054PR2
/ CURRENT APPLICATION NUMBER: US/09/621.976
/ CURRENT FILING DATE: 2000-07-21
```

```
/ NUMBER OF SEQ ID NOS: 19335
/ SOFTWARE: Patent.pm
/ SEQ ID NO 2230
/ LENGTH: 440
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: 32..253
/ US-09-621-976-2230

Query Match          3.4%; Score 16; DB 4; Length 440;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 391 CTGCTTTTGCTCTCTA 406
Db 240 CTGCTTTTGCTCTCTA 255

RESULT 14
US-09-134-001C-2724
/ Sequence 2724, Application US/09134001C
/ Patent No. 6380370
/ GENERAL INFORMATION:
/ APPLICANT: Lynn Doucette-Stamm et al
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
/ TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
/ FILE REFERENCE: GTC-007
/ CURRENT APPLICATION NUMBER: US/09/134,001C
/ CURRENT FILING DATE: 1998-08-13
/ PRIOR APPLICATION NUMBER: US 60/064,964
/ PRIOR FILING DATE: 1997-11-08
/ PRIOR APPLICATION NUMBER: US 60/055,779
/ PRIOR FILING DATE: 1997-08-14
/ NUMBER OF SEQ ID NOS: 5674
/ SEQ ID NO 2724
/ LENGTH: 813
/ TYPE: DNA
/ ORGANISM: Staphylococcus epidermidis
/ US-09-134-001C-2724

Query Match          3.4%; Score 16; DB 4; Length 813;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 24 TCTAATAAGCAATTAT 39
Db 274 TCTAATAAGCAATTAT 289

RESULT 15
US-08-956-171E-1111
/ Sequence 1111, Application US/08956171E
/ Patent No. 6593114
/ GENERAL INFORMATION:
/ APPLICANT: Charles Kunsch
/ APPLICANT: Gil H. Choi
/ APPLICANT: Patrick S. Dillon
/ APPLICANT: Craig A. Rosen
/ APPLICANT: Steven C. Barash
/ APPLICANT: Michael R. Fannon
/ TITLE OF INVENTION: Staphylococcus aureus Polynucleotides and Sequences
/ NUMBER OF SEQUENCES: 5256
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Human Genome Sciences, Inc.
/ STREET: 9410 Key West Avenue
/ CITY: Rockville
/ STATE: Maryland
/ COUNTRY: USA
/ ZIP: 20850
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
```

```
;
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/956,171E
; FILING DATE: 20-Oct-1997
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/009,861
; FILING DATE: January 5, 1996
; APPLICATION NUMBER: 08/781,986
; FILING DATE: January 3, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mark J. Hyman
; REGISTRATION NUMBER: 46,789
; REFERENCE/DOCKET NUMBER: PB248P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (240) 314-1224
; TELEFAX: (301) 309-8439
; INFORMATION FOR SEQ ID NO: 1111:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 825 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 1111:
US-08-956-171E-1111

Query Match          3.4%; Score 16; DB 4; Length 825;
Best Local Similarity 100.0%; Fred.No. 1.2e+02;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 418 TGTATTTCAATTGAA 433
   |||||||
Db 265 TGTATTTCAATTGAA 280
```

Search completed: July 31, 2004, 10:51:20  
Job time : 68 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2004, 08:13:50 ; Search time 65 Seconds  
(without alignments)  
4063.948 Million cell updates/sec

Title: US-10-020-540A-1  
Perfect score: 476  
Sequence: 1 aactatgctgacgaagat.....aacctccctgcttacaacc 476

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA: \*  
1: /cgn2\_6/ptodata/2/ina/5A COMB.seq.\*  
2: /cgn2\_6/ptodata/2/ina/5B COMB.seq.\*  
3: /cgn2\_6/ptodata/2/ina/6A COMB.seq.\*  
4: /cgn2\_6/ptodata/2/ina/6B COMB.seq.\*  
5: /cgn2\_6/ptodata/2/ina/PTUS COMB.seq.\*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	36.2	7.6	161652	4	US-09-497-855A-40
2	34.6	7.3	7218	1	US-08-232-463-14
3	34.4	7.2	1887	4	US-09-107-532A-2783
4	34	7.1	897	4	US-09-107-532A-72
5	32.4	6.8	1145	3	US-09-227-794-1
6	32.4	6.8	1677	4	US-09-252-991A-4734
7	32.4	6.8	2151	4	US-09-252-991A-4774
8	32	6.7	1923	4	US-09-328-352-3932
9	32	6.7	5170	4	US-09-677-046A-5
10	31.8	6.7	1215	3	US-08-642-274D-35
11	31.8	6.7	1215	3	US-08-952-014C-35
12	31.8	6.7	1482	4	US-09-328-352-2537
13	31.8	6.7	72604	4	US-09-268-992-7
14	31.8	6.7	72604	4	US-09-657-474-7
15	31.6	6.6	191	4	US-09-621-976-18332
16	31.6	6.6	1437	4	US-09-137-223A-4
17	31.6	6.6	14442	3	US-08-781-891-208
18	31.6	6.6	16442	4	US-09-618-186-208
19	31.6	6.6	58407	4	US-08-916-421B-2
20	31.4	6.6	870	4	US-09-107-532A-1977
21	31.4	6.6	2567	3	US-08-993-260-4
22	31.4	6.6	12482	4	US-09-512-563C-25
23	31.4	6.6	50000	4	US-09-146-053-4
24	31.2	6.6	5044	4	US-09-735-935-3
25	31.2	6.6	45716	4	US-08-965-048-5
26	31.2	6.6	45989	4	US-08-965-048-6
27	31	6.5	708	4	US-09-134-001C-1313

C	28	31	6.5	2664	4	US-09-328-352-353	Sequence 353, Appl
	29	30.8	6.5	201	4	US-09-134-000C-2031	Sequence 2031, Ap
	30	30.8	6.5	1434	4	US-09-134-000C-304	Sequence 304, App
	31	30.8	6.5	3475	4	US-08-924-829C-15	Sequence 15, Appl
C	32	30.6	6.4	600	4	US-09-621-976-10493	Sequence 10493, A
	33	30.6	6.4	2707	3	US-09-121-864-2	Sequence 2, Appli
	34	30.4	6.4	591	4	US-09-328-352-1319	Sequence 1319, Ap
	35	30.4	6.4	3942	4	US-09-601-198-50	Sequence 50, Appl
C	36	30.4	6.4	4977	1	US-08-030-096-7	Sequence 7, Appli
	37	30.4	6.4	10607	1	US-08-078-090-3	Sequence 3, Appli
	38	30.4	6.4	202001	4	US-09-734-674-3	Sequence 359, App
	39	30.2	6.3	643	4	US-08-956-171E-359	Sequence 2441, Ap
	40	30.2	6.3	1170	4	US-09-134-000C-2441	Sequence 4, Appli
C	41	30.2	6.3	1737	1	US-08-202-056-4	Sequence 3, Appli
	42	30.2	6.3	1737	1	US-08-076-093A-3	Sequence 3, Appli
C	43	30.2	6.3	1737	1	US-08-701-265-3	Sequence 3, Appli
	44	30.2	6.3	1737	2	US-08-284-586-3	Sequence 3, Appli
C	45	30.2	6.3	1737	2	US-08-805-478-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1

US-09-497-855A-40  
; Sequence 40, Application US/09497855A  
; Patent No. 6605432  
; GENERAL INFORMATION:

APPLICANT: Huang, Tim  
; TITLE OF INVENTION: HIGH-THROUGHPUT METHODS FOR DETECTING DNA METHYLATION  
; FILE REFERENCE: UM01523  
; CURRENT APPLICATION NUMBER: US/09/497,855A  
; CURRENT FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: 60/120,592  
; PRIOR FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: 60/118,760  
; PRIOR FILING DATE: 1999-02-05  
; NUMBER OF SEQ ID NOS: 54  
; SOFTWARE: Patent in version 3.0  
; SEQ ID NO 40  
; LENGTH: 161652  
; TYPE: DNA  
; ORGANISM: Homo sapiens;  
US-09-497-855A-40

Query Match 7.6%; Score 36.2; DB 4; Length 161652;  
Best Local Similarity 46.9%; Pred. No. 2.7;  
Matches 113; Conservative 0; Mismatches 128; Indels 0; Gaps 0;

QY	152	CACCAGAACCATCTTTGCGACGATGTGAAGCAGCCAGAGTGTCCACAGACGCACT	211
Db	21744	CACTCTCACCCCTTCTATTCAACATAGGACTTAAGTCTTAGCATAGGTAGAAAAGA	21803
QY	212	CAGAAAGGACATCTTTACCGACACAGAAAAGAACACACAGCTCATCCACATGT	271
Db	21804	AACAAAGGCATCTAAATCAGAAAAGGAGTAATGATCTCTGTTTTAGATGACTT	21863
QY	272	AGACGTCTCTTATCGCTCGGCTGAGATAGACTGACCCGCGCAGCACTAAAGAGAA	331
Db	21864	GATCTTACATGAGAAAATCCCTAAGACACCCACAAAATTTGTTAGACTTAATGA	21923
QY	332	ATAATCAAGTGGTCTCTAGCTCCACTTTAGCTTTTAAATTAATTTGTTTATTTC	391
Db	21924	ATTCACTCAAGTTCGAGATACAAATCAACATATAAAATCAGTTGTTATTTGTACAC	21983
QY	392	T 392	
Db	21984	T 21984	

RESULT 2

US-08-232-463-14/c  
; Sequence 14, Application US/08232463

```
Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHEIFLINGER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: pTZgpt-F1s
US-08-232-463-14

Query Match 7.3%; Score 34.6; DB 1; Length 7218;
Best Local Similarity 2.0%; Pred.No.1.9;
Matches 7; Conservative 192; Mismatches 146; Indels 0; Gaps 0;

QY 10 GATGACAAGATAATCTCTAATAGCAATATTTCAGAAATTAATCAAGAGAAAGAAATTAATA 69
DB 1399 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1340
QY 70 ACTCTTTTCAGAAATGAAAGCCCGCTTTTCAAGAGTGGCAGCTAGCTATCATCTGAAAGACA 129
DB 1339 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1280
QY 130 GCAAGACAATGGTGTCTGATGACCAAGAACACATCTTTGCACGACATGTGAGACGCC 189
DB 1279 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1220
QY 190 AGAGTGTGTCACAGACCACTCAAGAAAGCATCTTTCTACCGACACAGAAAAAGACAAC 249
DB 1219 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1160
QY 250 CACAGCTCATCATCAACATCTAGACTGCTGTTATGCTGGCTGAGATAAGACTGACC 309
DB 1159 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1100
QY 310 CCAGCCGACACTAAGAGAAATTAATGCAAGTGGTCTAGCTCC 354
DB 1099 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1055

RESULT 3
US-09-107-532A-2783/c
; Sequence 2783, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
NUMBER OF SEQUENCES: 7310
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02154
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: PC
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,532A
FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Ariniello, Pamela Deneké
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 2783:
SEQUENCE CHARACTERISTICS:
LENGTH: 1887 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium
FEATURE:
NAME/KEY: misc feature
LOCATION: (3) LOCATION 1....1887
SEQUENCE DESCRIPTION: SEQ ID NO: 2783:
US-09-107-532A-2783

Query Match 7.2%; Score 34.4; DB 4; Length 1887;
Best Local Similarity 57.4%; Pred.No.1.1;
Matches 62; Conservative 0; Mismatches 46; Indels 0; Gaps 0;

QY 323 AAAGAAGAAATTAATGCAAGTGGTCTAGCTCCACTTTAGCTTTAATATTAATTTGTTTCA 382
DB 511 AAACAAAATAAAATAAAGTAGTCGTAATTTAGCTCATCTCTCTATTATTCTTAGTCCT 452
QY 383 ATTATTCTCTGCTTTTGTCTCTATATAAAGAGCTTGTATTTCATTT 430
DB 451 TATTGGTGTGCTGCTAATCTTTATCAATCAAAATCATTCATCTCTTT 404

RESULT 4
US-09-107-532A-72
; Sequence 72, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
```

```

; FILE REFERENCE: 0945
; CURRENT APPLICATION NUMBER: US/09/227,794A
; CURRENT FILING DATE: 1999-01-08
; EARLIER APPLICATION NUMBER: US 09/227,794
; EARLIER FILING DATE: 1999-01-08
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 1145
; TYPE: DNA
; ORGANISM: Glycine max
US-09-227-794-1

Query Match          6.8%; Score 32.4; DB 3; Length 1145;
Best Local Similarity 52.2%; Pred. No. 3.6;
Matches 72; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

QY      296 AGATGAAGACTGACCCAGCCAGCAGCACTAAAGAGAAATAATGCAAGTGGTCTCTAGCTCCA 355
        |||||+|||||
DB      459 AGATCAATTTGAGTGAATCATCTTCTGTAAACAATAATGAAGTGGTTAAAGCACAT 400
        |||||+|||||
QY      356 CTTTAGCTTTAATAATTATGTTTCATTATTATTTCTCTGCTTTTGTCTCTCTATATAAGAG 415
        |||||+|||||
DB      399 TTTTGTATTTTTTTACTTTTTTTTGTCTCTTCATTTTAGTTTTTAAAAATAAAGTAAATAT 340
        |||||+|||||
QY      416 CTGTGATTTTCATTTGAA 433
DB      339 AAGTTCTGTCTTTTGTA 322

RESULT 6
US-09-252-991A-4734/c
; Sequence 4734, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107195.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 4734
; LENGTH: 1677
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-4734

```

```

; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-4734

Query Match          6.8%; Score 32.4; DB 4; Length 1677;
Best Local Similarity 52.2%; Pred. No. 4; 4;
Matches 72; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

Qy      216 AAAGGATCTTTACCGCACAGAAAAGACACACAGCTCATCTCAACATGTAGAC 275
Db      1539 AAGGGATGACCTACCGCGAGAGAAATCTTCGGAGCTCAAGGACCTGGCCGAGAG 1480

Qy      276 TGTGGTTATGCGTCGGCTGAAGATAAGATGACCCAGCGCCAGCATTAAGAAGAAATAA 335
Db      1479 TGGCGTTCCAGCATGTGCGAGGCTGGCGGCCGAAGCCACGAAGAGCTGTATGAACAAAGTAC 1420

Qy      336 TGCAAGTGGTCCTAGCTC 353
Db      1419 CTCGAAGAGGGCGAGCTC 1402

```

RESULT 7  
US-03-352-991A-4774  
; Sequence 4774, Application US/09252991A  
; Patent No. 6551795

Db 1419 CTCGAGAGGGCGAGCTC 1402





APPLICANT: Shiloh, Yosef  
TITLE OF INVENTION: ATAXIA-TELANGIECTASIA GENE AND ITS  
TITLE OF INVENTION: GENOMIC ORGANIZATION  
NUMBER OF SEQUENCES: 91  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Kohn & Associates  
STREET: 30500 No. 6265158thwestern Hwy., Suite 410  
CITY: Farmington Hills  
STATE: Michigan  
COUNTRY: U.S.  
ZIP: 48334

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA: US/08/952,014C  
APPLICATION NUMBER: US/08/952,014C  
FILING DATE:

CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Kohn, Kenneth I.  
REGISTRATION NUMBER: 30,995  
REFERENCE/DOCKET NUMBER: 2290.00028  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 810-539-5050  
TELEFAX: 810-539-5055

INFORMATION FOR SEQ ID NO: 35:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1215 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
US-08-952-014C-35

Query Match 6.7%; Score 31.8; DB 3; Length 1215;  
Best Local Similarity 50.3%; Pred. No. 5.7;  
Matches 78; Conservative 0; Mismatches 77; Indels 0; Gaps 0;  
QY 320 ACTAAGAAGAAATAATGCAAGTGGCTCCTAGCTCCACTTTAGCTTTAATAATTATGTTTC 379  
Db 277 AGTTAAAGCAATTTTAGAATCTTCCAAAAGCCTGTAAGTATACATGATGAGTTTA 336  
QY 380 ATTATTAATCTGCTTTTCTCTCTATATAAGAGCTTGATTTCAATTTGAGGCAGA 439  
Db 337 ATAATAGAACATTCCTCTCTTTTGTAGCTAAAGAACTTTGTAATACATCTTAAAGAGGA 396  
QY 440 GGCGAACACACACAGAACCTCCCTGCTTACAAA 474  
Db 397 AAGTAACAAGAAAGAAATTTATCTCATATTA 431

RESULT 12  
US-09-328-352-2537  
; Sequence 2537, Application US/09328352  
; Patent No. 6562958  
; GENERAL INFORMATION:  
; APPLICANT: Gary L. Breton et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: GTC99-03PA  
; CURRENT APPLICATION NUMBER: US/09/328,352  
; CURRENT FILING DATE: 1999-06-04  
; NUMBER OF SEQ ID NOS: 8252  
; SEQ ID NO 2537  
; LENGTH: 1482  
; TYPE: DNA  
; ORGANISM: Acinetobacter baumannii  
US-09-328-352-2537

Query Match 6.7%; Score 31.8; DB 4; Length 1482;  
Best Local Similarity 52.7%; Pred. No. 6.3;

Matches 69; Conservative 0; Mismatches 62; Indels 0; Gaps 0;  
QY 3 CTATGCTGATCACAAGATTAATCTTAATAAGCAATTTATTCAGATTAATCAAGGAGAAAGA 62  
Db 120 CTTTACCATGAAAAAATTTGTTGAATTTACTTTAAGATCAACATGATCTGAATTTGGTAAC 179  
QY 63 ATTAATAACTCTTTTCAGAAATATGAAGCCCGCTTTTACAAGTGGCCAGCTAGCTATCACTGA 122  
Db 180 AGTCAAAACATTTAAAGTTATTATATACCGATATGCAAAAGTGAATCCAGCTATGTCTAA 239  
QY 123 AAAGACAGCAA 133  
Db 240 AAATACTGCTA 250

RESULT 13  
US-09-268-992-7/C  
; Sequence 7, Application US/09268992  
; Patent No. 6342351  
; GENERAL INFORMATION:  
; APPLICANT: Chen, H.  
; APPLICANT: Freimer, N.  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING  
; TITLE OF INVENTION: AND TREATING CHROMOSOME-18p RELATED DISORDERS  
; FILE REFERENCE: 7853-138  
; CURRENT APPLICATION NUMBER: US/09/268,992  
; CURRENT FILING DATE: 1999-03-16  
; EARLIER APPLICATION NUMBER: 09/236,134  
; EARLIER FILING DATE: 1999-01-22  
; EARLIER APPLICATION NUMBER: 60/106,056  
; EARLIER FILING DATE: 1998-10-28  
; EARLIER APPLICATION NUMBER: 60/088,312  
; EARLIER FILING DATE: 1998-06-05  
; EARLIER APPLICATION NUMBER: 60/078,044  
; EARLIER FILING DATE: 1998-03-16  
; NUMBER OF SEQ ID NOS: 84  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 7  
; LENGTH: 72604  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: modified base  
; LOCATION: all n positions  
; OTHER INFORMATION: n=a, c, g, or t  
US-09-268-992-7

Query Match 6.7%; Score 31.8; DB 4; Length 72604;  
Best Local Similarity 49.7%; Pred. No. 40;  
Matches 81; Conservative 0; Mismatches 82; Indels 0; Gaps 0;  
QY 278 TCGTTATGCTGGCTGAAGATAGACTGACCCAGGCCAGCACTAAGAGAGAAATATG 337  
Db 23019 TAGTTCCTTCCTGGTCATAGATAAGAACACAGACCTAACTGAGCTAAGGACCAAAAAAT 22960  
QY 338 CAAAGTGGCTCTAGCTCCACTTTAGCTTTAATAATTTATTTCAATTTATTTCTCTGCTTT 397  
Db 22959 CTGCATCACTGTTGCAAGTTTACTTTAGATTTTAACTTATAGTTTATGTTAAATTAC 22900  
QY 398 TCGCTCTATATAAGACCTTGTTATTTTCATTTGAGGCAGAG 440  
Db 22899 TGTTCATTTTGAAGAGGATACATTTTGGTTATTAATAAAGTG 22857

RESULT 14  
US-09-657-474-7/C  
; Sequence 7, Application US/09657474  
; Patent No. 6399762  
; GENERAL INFORMATION:  
; APPLICANT: Chen, H.  
; APPLICANT: Freimer, N.  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DIAGNOSING  
; TITLE OF INVENTION: AND TREATING CHROMOSOME-18p RELATED DISORDERS

```

; FILE REFERENCE: 7853-138
; CURRENT APPLICATION NUMBER: US/09/657,474
; CURRENT FILING DATE: 2000-09-07
; PRIOR APPLICATION NUMBER: 09/268,992
; PRIOR FILING DATE: 1999-03-16
; PRIOR APPLICATION NUMBER: 09/236,134
; PRIOR FILING DATE: 1999-01-22
; PRIOR APPLICATION NUMBER: 60/106,056
; PRIOR FILING DATE: 1998-10-28
; PRIOR APPLICATION NUMBER: 60/088,312
; PRIOR FILING DATE: 1998-06-05
; PRIOR APPLICATION NUMBER: 60/078,044
; PRIOR FILING DATE: 1998-03-16
; NUMBER OF SEQ ID NOS: 84
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 72604
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: modified base
; LOCATION: all n positions
; OTHER INFORMATION: n=a, c, g, or t
US-09-657-474-7

```

	Query Match	6.7%;	Score 31.8;	DB 4;	Length 72604;
	Best Local Similarity	49.7%;	Pred. No. 40;		
	Matches	81;	Conservative 0;	Mismatches 82;	Indels 0; Gaps 0;
Qy	278	TCGTTATGGCTGGCTGAGATAAGACTGACCCAGGCCAGCACCTAAGAGAGAAATATG	337		
Db	23019	TAGTTCTTCCTGGTCATAGATAAGAACACAGACCTAACTGAGCTAAGGACGACAAAATT	22960		
Qy	338	CAAGTGGCTCCTAGCTCCACTTTAGCTTTAAATAATTATGTTTCATTATTATTCTCTGCTTT	397		
Db	22959	CTGCATCACGTGTGCAAGTTTACTTTAGATTATTAAGTTTCATAGTTTATGTTAAATTAC	22900		
Qy	398	TGCTCTCTATATAAAGAGCTTGTATTTTCATTGTAAGGCAGAG	440		
Db	22899	TGTTTTTCATTTCGAAAGGATACATTTCGTTTATTAATAAAGTG	22857		

```

RESULT 15
US-09-621-976-18332/c
; Sequence 18332, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 18332
; LENGTH: 191
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-18332

Query Match          6.6%; Score 31.6; DB 4; Length 191;
Best Local Similarity 60.5%; Pred. No. 2.7;
Matches 52; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

QY      357  TTTAGCTTTAATAATTAAGTTTCATTATTATTCCTGCTCTCTATATAAGAGC 416
          |||
          198  TTTTNTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTGGAGTAATAACTGA 129

Db

QY      417  TGTATTTTTCATTTCGAGGCGAGGC 442

Db      128  TTTTAACTTTATTTTACTGCAGAGAC 103

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2004, 08:12:51 ; Search time 335 Seconds  
(without alignments)  
6966.850 Million cell updates/sec

Title: US-10-020-540A-1

Perfect score: 476

Sequence: 1 aactatgctgatacaagat.....aacctccctgcttacaacc 476

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3222919 seqs, 2451570024 residues

Total number of hits satisfying chosen parameters: 6445838

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:\*

- 1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*
- 4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq:\*
- 10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq:\*
- 12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*
- 13: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*
- 14: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq:\*
- 15: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq:\*
- 16: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq:\*
- 17: /cgn2\_6/ptodata/2/pubpna/US10C\_NEW\_PUB.seq:\*
- 18: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*
- 19: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	476	100.0	476	14	US-10-020-540A-1
2	476	100.0	518	14	US-10-020-540A-2
3	369	77.5	651	14	US-10-020-540A-3
4	369	77.5	701	14	US-10-020-540A-4
5	40.2	8.4	539	13	US-10-027-632-72822
6	40.2	8.4	539	13	US-10-027-632-312801
7	40.2	8.4	539	16	US-10-027-632-72822
8	40.2	8.4	539	16	US-10-027-632-312801
9	40.2	8.4	547	13	US-10-027-632-49899
10	40.2	8.4	547	16	US-10-027-632-49899
11	37	7.8	90798	17	US-10-318-819A-4
12	37	7.8	3673778	15	US-10-313-841-2
13	36.8	7.7	6012	13	US-10-221-714A-449
14	36.6	7.7	17211	15	US-10-311-455-628

C	15	36.4	7.6	1141	13	US-10-424-599-93424	Sequence 93424, A
	16	36.2	7.6	161652	15	US-10-081-327-40	Sequence 40, Appl
	17	35.8	7.5	379	13	US-10-424-599-132308	Sequence 132308, A
	18	35.8	7.5	422	9	US-09-783-590-4784	Sequence 4784, Ap
C	19	35.6	7.5	586	13	US-10-027-632-105291	Sequence 105291, A
C	20	35.6	7.5	586	16	US-10-027-632-105291	Sequence 105291, A
C	21	35.4	7.4	593	13	US-10-027-632-214128	Sequence 214128, A
	22	35.4	7.4	593	16	US-10-027-632-214128	Sequence 214128, A
	23	35	7.4	458	9	US-09-764-853-272	Sequence 272, App
	24	35	7.4	458	9	US-09-764-860-86	Sequence 86, Appl
	25	35	7.4	458	15	US-10-091-438-53	Sequence 53, Appl
	26	35	7.4	458	15	US-10-074-095-86	Sequence 86, Appl
	27	35	7.4	458	16	US-10-212-872-86	Sequence 86, Appl
	28	35	7.4	494	9	US-09-783-590-4851	Sequence 4851, Ap
	29	35	7.4	502	9	US-09-764-853-15	Sequence 15, Appl
	30	35	7.4	504	9	US-09-764-860-770	Sequence 770, App
	31	35	7.4	504	15	US-10-091-438-265	Sequence 265, App
	32	35	7.4	504	15	US-10-074-095-770	Sequence 770, App
	33	35	7.4	504	16	US-10-212-872-770	Sequence 770, App
	34	35	7.4	573	13	US-10-027-632-270337	Sequence 270337, A
	35	35	7.4	729	13	US-10-027-632-15321	Sequence 15321, A
C	36	35	7.4	729	16	US-10-027-632-15321	Sequence 15321, A
C	37	35	7.4	1344	13	US-10-425-114-6401	Sequence 6401, Ap
	38	35	7.4	2693	13	US-10-424-599-97948	Sequence 97948, A
C	39	35	7.4	655	16	US-10-341-961A-245	Sequence 245, App
C	40	34.8	7.3	6314	9	US-09-764-887-473	Sequence 473, App
C	41	34.6	7.3	6314	15	US-10-073-961-473	Sequence 473, App
C	42	34.6	7.3	9289	9	US-09-764-887-471	Sequence 471, App
C	43	34.6	7.3	9289	15	US-10-073-961-471	Sequence 471, App
C	44	34.6	7.3	17865	9	US-09-764-887-472	Sequence 472, App
C	45	34.6	7.3	17865	15	US-09-764-887-472	Sequence 472, App

ALIGNMENTS

RESULT 1

US-10-020-540A-1  
; Sequence 1, Application US/10020540A  
; Publication No. US20020182593A1  
; GENERAL INFORMATION:  
; APPLICANT: Wu, Gusui  
; TITLE OF INVENTION: Strawberry Vein Banding Virus Promoter  
; TITLE OF INVENTION: (SVBV) Promoter  
; FILE REFERENCE: 0173.210us  
; CURRENT APPLICATION NUMBER: US/10/020,540A  
; PRIOR FILING DATE: 2001-10-30  
; PRIOR APPLICATION NUMBER: 60/245,354  
; PRIOR FILING DATE: 2000-11-01  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 476  
; TYPE: DNA  
; ORGANISM: Strawberry Vein Banding Virus (SVBV) Strain E3  
US-10-020-540A-1

Query Match 100.0%; Score 476; DB 14; Length 476;  
Best Local Similarity 100.0%; Pred. No. 1e-127;  
Matches 476; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	AACATGCTGATGACAAGATAATTTCTAATAGCANTATTTCAGATTATTAATCAAGGAGAA	60
DB	1	AACATGCTGATGACAAGATAATTTCTAATAGCANTATTTCAGATTATTAATCAAGGAGAA	60
QY	61	GAATTAATAACTCTTTTCAGATATGAAAGCCGCTTTACAAGTGGCCAGCTAGCTACT	120
DB	61	GAATTAATAACTCTTTTCAGATATGAAAGCCGCTTTACAAGTGGCCAGCTAGCTACT	120
QY	121	GAAGAGACAGCAAGACATGGTGTCTCGATGACAGAACACATCTTTGACAGAGATGT	180
DB	121	GAAGAGACAGCAAGACATGGTGTCTCGATGACAGAACACATCTTTGACAGAGATGT	180

QY 181 GAAGCAGCAGAGTGGTCCACAGAGCGACTCAGAAAAGGCATCTTCTACCGACACAGAA 240  
DB 181 GAAGCAGCAGAGTGGTCCACAGAGCGACTCAGAAAAGGCATCTTCTACCGACACAGAA 240  
QY 241 AAAGCAACACACAGCTCATCTCAACATGTAGACTGTCTGTTATGCTGCGCTCGAGATA 300  
DB 241 AAAGCAACACACAGCTCATCTCAACATGTAGACTGTCTGTTATGCTGCGCTCGAGATA 300  
QY 301 AGACTGACCCAGCCAGCCAGCAGCTAAAGAGAAATATGCAAGTGGTCTAGCTCCACTTTA 360  
DB 301 AGACTGACCCAGCCAGCCAGCAGCTAAAGAGAAATATGCAAGTGGTCTAGCTCCACTTTA 360  
QY 361 GCTTTAATAATATGTTTCAATATTATTTCTGCTTTTCTCTCTATATAAAGAGCTTGT 420  
DB 361 GCTTTAATAATATGTTTCAATATTATTTCTGCTTTTCTCTCTATATAAAGAGCTTGT 420  
QY 421 ATTTTCATTTGAAGCAGAGCGGAACACACACAGAAACCTCCCTGCTTACAAACC 476  
DB 421 ATTTTCATTTGAAGCAGAGCGGAACACACACAGAAACCTCCCTGCTTACAAACC 476

RESULT 2  
US-10-020-540A-2  
; Sequence 2, Application US/10020540A  
; Publication No. US20020182593A1  
; GENERAL INFORMATION:  
; APPLICANT: Wu, Gusui  
; TITLE OF INVENTION: Strawberry Vein Banding Virus Promoter  
; FILE REFERENCE: 0173.210us  
; CURRENT APPLICATION NUMBER: US/10/020,540A  
; CURRENT FILING DATE: 2001-10-30  
; PRIOR APPLICATION NUMBER: 60/245,354  
; PRIOR FILING DATE: 2000-11-01  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 518  
; TYPE: DNA  
; ORGANISM: Strawberry Vein Banding Virus (SVBV) Strain E3  
US-10-020-540A-2

Query Match 100.0%; Score 476; DB 14; Length 518;  
Best Local Similarity 100.0%; Pred. No. 1.1e-127;  
Matches 476; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AACTATGCTGATGACAAGATAATTTCTAATAAGCAATATTTCAGAAATTAATCAAGAGAAA 60  
DB 15 AACTATGCTGATGACAAGATAATTTCTAATAAGCAATATTTCAGAAATTAATCAAGAGAAA 74  
QY 61 GAATTAATAACTCTTTTCAAGATATGAAGCCGCTTTTACAAGTGGCCAGCTAGCTATCACT 120  
DB 75 GAATTAATAACTCTTTTCAAGATATGAAGCCGCTTTTACAAGTGGCCAGCTAGCTATCACT 134  
QY 121 GAAAAGACAGCAGCAAGCAATGGTGTCTCGATGACACAGAAACCAATCTTTTGACGAGATGT 180  
DB 135 GAAAAGACAGCAGCAAGCAATGGTGTCTCGATGACACAGAAACCAATCTTTTGACGAGATGT 194  
QY 181 GAAGCAGCAGAGTGGTCCACAGAGCGACTCAGAAAAGGCATCTTCTACCGACACAGAA 240  
DB 195 GAAGCAGCAGAGTGGTCCACAGAGCGACTCAGAAAAGGCATCTTCTACCGACACAGAA 254  
QY 241 AAAGCAACACACAGCTCATCTCAACATGTAGACTGTCTGTTATGCTGCGCTCGAGATA 300  
DB 255 AAAGCAACACACAGCTCATCTCAACATGTAGACTGTCTGTTATGCTGCGCTCGAGATA 314  
QY 301 AGACTGACCCAGCCAGCCAGCAGCTAAAGAGAAATATGCAAGTGGTCTAGCTCCACTTTA 360  
DB 315 AGACTGACCCAGCCAGCCAGCAGCTAAAGAGAAATATGCAAGTGGTCTAGCTCCACTTTA 374  
QY 361 GCTTTAATAATATGTTTCAATATTATTTCTGCTTTTCTCTATATAAAGAGCTTGT 420

DB 375 GCTTTAATAATATGTTTCAATATTATTTCTGCTTTTCTCTATATAAAGAGCTTGT 434  
QY 421 ATTTTCATTTGAAGCAGAGCGGAACACACACAGAAACCTCCCTGCTTACAAACC 476  
DB 435 ATTTTCATTTGAAGCAGAGCGGAACACACACAGAAACCTCCCTGCTTACAAACC 490

RESULT 3  
US-10-020-540A-3  
; Sequence 3, Application US/10020540A  
; Publication No. US20020182593A1  
; GENERAL INFORMATION:  
; APPLICANT: Wu, Gusui  
; TITLE OF INVENTION: Strawberry Vein Banding Virus Promoter  
; FILE REFERENCE: 0173.210us  
; CURRENT APPLICATION NUMBER: US/10/020,540A  
; CURRENT FILING DATE: 2001-10-30  
; PRIOR APPLICATION NUMBER: 60/245,354  
; PRIOR FILING DATE: 2000-11-01  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 651  
; TYPE: DNA  
; ORGANISM: Strawberry Vein Banding Virus (SVBV) Strain E3  
US-10-020-540A-3

Query Match 77.5%; Score 369; DB 14; Length 651;  
Best Local Similarity 100.0%; Pred. No. 1.5e-96;  
Matches 369; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 108 GCTAGCTATCACTGAAAAGACACAGCAAGCAATGGTGTCTCGATGACACAGAACCACTCT 167  
DB 278 GCTAGCTATCACTGAAAAGACACAGCAAGCAATGGTGTCTCGATGACACAGAACCACTCT 337  
QY 168 TTGCAGCAGATGTGAAGCAGCAGAGTGTCTCACAAGACGCGACTCAGAAAAGGCATCTTC 227  
DB 338 TTGCAGCAGATGTGAAGCAGCAGAGTGTCTCACAAGACGCGACTCAGAAAAGGCATCTTC 397  
QY 228 TACCGACACAGAAAAGACACACAGCTCATCTCAACATGTAGACTGTGTTATGG 287  
DB 398 TACCGACACAGAAAAGACACACAGCTCATCTCAACATGTAGACTGTGTTATGG 457  
QY 288 TCGGCTGAAGATGAAGTGTGACCCAGCCAGCAGCTAAAGAGAAATATGCAAGTGGTCC 347  
DB 458 TCGGCTGAAGATGAAGTGTGACCCAGCCAGCAGCTAAAGAGAAATATGCAAGTGGTCC 517  
QY 348 TAGCTCCACTTTAGCTTTTAATAATATTATGTTTCATTATTATTTCTGCTTTTCTCTAT 407  
DB 518 TAGCTCCACTTTAGCTTTTAATAATATTATGTTTCATTATTATTTCTGCTTTTCTCTAT 577  
QY 408 ATAAAGAGCTGTGTTTTCATTGAGGCGAGGGGGAACACACACAGAAACCTCCCTGC 467  
DB 578 ATAAAGAGCTGTGTTTTCATTGAGGCGAGGGGGAACACACACAGAAACCTCCCTGC 637  
QY 468 TTACAAACC 476  
DB 638 TTACAAACC 646

RESULT 4  
US-10-020-540A-4  
; Sequence 4, Application US/10020540A  
; Publication No. US20020182593A1  
; GENERAL INFORMATION:  
; APPLICANT: Wu, Gusui  
; TITLE OF INVENTION: Strawberry Vein Banding Virus Promoter  
; FILE REFERENCE: 0173.210us

```
; CURRENT APPLICATION NUMBER: US/10/020,540A
; CURRENT FILING DATE: 2001-10-30
; PRIOR APPLICATION NUMBER: 60/245,354
; PRIOR FILING DATE: 2000-11-01
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 701
; TYPE: DNA
; ORGANISM: Strawberry Vein Banding Virus (SVBV) Strain E3
US-10-020-540A-4

Query Match      77.5%; Score 369; DB 14; Length 701;
Best Local Similarity 100.0%; Pred. No. 1.5e-96;
Matches 369; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 108 GCTAGCTATCACTGAAAGACAGACAAATGGTGTCTCGATGCCACCAAGACCAATCT 167
DB 108 GCTAGCTATCACTGAAAGACAGACAAATGGTGTCTCGATGCCACCAAGACCAATCT 167
QY 168 TTGCAGCAGATGTGAAGCAGCCAGAGTGTCTCAAGAGCCACTCAGAAAAGGCATCTTC 227
DB 168 TTGCAGCAGATGTGAAGCAGCCAGAGTGTCTCAAGAGCCACTCAGAAAAGGCATCTTC 227
QY 379 TTGCAGCAGATGTGAAGCAGCCAGAGTGTCTCAAGAGCCACTCAGAAAAGGCATCTTC 438
DB 379 TTGCAGCAGATGTGAAGCAGCCAGAGTGTCTCAAGAGCCACTCAGAAAAGGCATCTTC 438
QY 228 TACCGACACAGAAAAGACACACAGCTCATCATCAACATGTAGACTGTCTGTATGCG 287
DB 228 TACCGACACAGAAAAGACACACAGCTCATCATCAACATGTAGACTGTCTGTATGCG 287
QY 439 TACCGACACAGAAAAGACACACAGCTCATCATCAACATGTAGACTGTCTGTATGCG 498
DB 439 TACCGACACAGAAAAGACACACAGCTCATCATCAACATGTAGACTGTCTGTATGCG 498
QY 288 TCGGCTGAAGATAAGACTGACCCAGCCAGCAGCTAAAGAGAAATAATGCAAGTGGTCC 347
DB 288 TCGGCTGAAGATAAGACTGACCCAGCCAGCAGCTAAAGAGAAATAATGCAAGTGGTCC 347
QY 499 TCGGCTGAAGATAAGACTGACCCAGCCAGCAGCTAAAGAGAAATAATGCAAGTGGTCC 558
DB 499 TCGGCTGAAGATAAGACTGACCCAGCCAGCAGCTAAAGAGAAATAATGCAAGTGGTCC 558
QY 348 TAGCTCCACTTTAGCTTTAATAATTATTTTCATTTATTTCTCTCTCTCTCTCTAT 407
DB 348 TAGCTCCACTTTAGCTTTAATAATTATTTTCATTTATTTCTCTCTCTCTCTCTAT 407
QY 559 TAGCTCCACTTTAGCTTTAATAATTATTTTCATTTATTTCTCTCTCTCTCTCTAT 618
DB 559 TAGCTCCACTTTAGCTTTAATAATTATTTTCATTTATTTCTCTCTCTCTCTCTAT 618
QY 408 ATAAAGAGCTTGTATTTTCATTTTGAAGCGAGCGCAACACACACAGAACTCCCTGCG 467
DB 408 ATAAAGAGCTTGTATTTTCATTTTGAAGCGAGCGCAACACACACAGAACTCCCTGCG 467
QY 619 ATAAAGAGCTTGTATTTTCATTTTGAAGCGAGCGCAACACACACAGAACTCCCTGCG 678
DB 619 ATAAAGAGCTTGTATTTTCATTTTGAAGCGAGCGCAACACACACAGAACTCCCTGCG 678
QY 468 TTCAAAACC 476
DB 468 TTCAAAACC 476
QY 679 TTCAAAACC 687
DB 679 TTCAAAACC 687

RESULT 5
US-10-027-632-72822/c
; Sequence 72822, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 312801
; LENGTH: 539
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-72822

Query Match      8.4%; Score 40.2; DB 13; Length 539;
Best Local Similarity 53.5%; Pred. No. 0.41;
Matches 84; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

QY 233 ACACAGAAAAGACAAACACAGCTCATCATCAACATGTAGACTGTCTGTATGCGTCGGC 292
DB 233 ACACAGAAAAGACAAACACAGCTCATCATCAACATGTAGACTGTCTGTATGCGTCGGC 292
QY 423 AAACAGAAAAGAAACCTCAGGCCAATATCCATGATGAACATTCACACAAAATCCCTCA 364
DB 423 AAACAGAAAAGAAACCTCAGGCCAATATCCATGATGAACATTCACACAAAATCCCTCA 364
QY 293 TGAAGATAAGACTGACCCAGCCAGCAGCTAAAGAGAAATAATGCAAGTGGTCTCTAGCT 352
DB 293 TGAAGATAAGACTGACCCAGCCAGCAGCTAAAGAGAAATAATGCAAGTGGTCTCTAGCT 352
QY 363 TTAAATAATAGTCAACCAATCCAGCAGCAATCAGAAAGTTTATCTACCAATAATCAAG 304
DB 363 TTAAATAATAGTCAACCAATCCAGCAGCAATCAGAAAGTTTATCTACCAATAATCAAG 304
QY 353 CCACCTTAGCTTTAATAATTATTTTCATTTATTTTAC 389
DB 353 CCACCTTAGCTTTAATAATTATTTTCATTTATTTTAC 389
QY 303 TTGGCTTCACCTCCAGGCAATTTATTTTAC 267
DB 303 TTGGCTTCACCTCCAGGCAATTTATTTTAC 267

RESULT 6
US-10-027-632-312801/c
; Sequence 312801, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 312801
; LENGTH: 539
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-312801

Query Match      8.4%; Score 40.2; DB 13; Length 539;
Best Local Similarity 53.5%; Pred. No. 0.41;
Matches 84; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

QY 233 ACACAGAAAAGACAAACACAGCTCATCATCAACATGTAGACTGTCTGTATGCGTCGGC 292
DB 233 ACACAGAAAAGACAAACACAGCTCATCATCAACATGTAGACTGTCTGTATGCGTCGGC 292
QY 423 AAACAGAAAAGAAACCTCAGGCCAATATCCATGATGAACATTCACACAAAATCCCTCA 364
DB 423 AAACAGAAAAGAAACCTCAGGCCAATATCCATGATGAACATTCACACAAAATCCCTCA 364
QY 293 TGAAGATAAGACTGACCCAGCCAGCAGCTAAAGAGAAATAATGCAAGTGGTCTCTAGCT 352
DB 293 TGAAGATAAGACTGACCCAGCCAGCAGCTAAAGAGAAATAATGCAAGTGGTCTCTAGCT 352
QY 363 TTAAATAATAGTCAACCAATCCAGCAGCAATCAGAAAGTTTATCTACCAATAATCAAG 304
DB 363 TTAAATAATAGTCAACCAATCCAGCAGCAATCAGAAAGTTTATCTACCAATAATCAAG 304
QY 353 CCACCTTAGCTTTAATAATTATTTTCATTTATTTTAC 389
DB 353 CCACCTTAGCTTTAATAATTATTTTCATTTATTTTAC 389
QY 303 TTGGCTTCACCTCCAGGCAATTTATTTTAC 267
DB 303 TTGGCTTCACCTCCAGGCAATTTATTTTAC 267

RESULT 7
```

```
US-10-027-632-72822/c
; Sequence 72822, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 72822
; LENGTH: 539
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-72822

Query Match      8.4%; Score 40.2; DB 16; Length 539;
Best Local Similarity 53.5%; Pred. No. 0.41;
Matches 84; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

QY 233 ACACAGAAAAGACACACAGCTCATCCACATGTAGACTGCTGTTATGCGTCGGC 292
Db 423 AAACAAAAAAGAAACCTCAGGCCAATATCCATGATGAACATTGACACAAAAATCCTCA 364
QY 293 TGAAGATAAGACTGACCCAGCCAGCAGCTAAAGAAAGAAATAATGCAAGTGTCTAGCT 352
Db 363 TTAATAATAAGTCAACCAATCCAGCAGCAATCAGAAAGTTTATCTACCATATCAAG 304
QY 353 CCACCTTAGCTTTAATAATTATGTTTCATTATTTC 389
Db 303 TTGGCTTCACTCCAGGCATTTTATTCAGTTTTTTAC 267

RESULT 8
US-10-027-632-312801/c
; Sequence 312801, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 312801
; LENGTH: 539
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-312801/c

Query Match      8.4%; Score 40.2; DB 16; Length 539;
Best Local Similarity 53.5%; Pred. No. 0.41;
Matches 84; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

QY 233 ACACAGAAAAGACACACAGCTCATCCACATGTAGACTGCTGTTATGCGTCGGC 292
Db 423 AAACAAAAAAGAAACCTCAGGCCAATATCCATGATGAACATTGACACAAAAATCCTCA 364
QY 293 TGAAGATAAGACTGACCCAGCCAGCAGCTAAAGAAAGAAATAATGCAAGTGTCTAGCT 352
Db 363 TTAATAATAAGTCAACCAATCCAGCAGCAATCAGAAAGTTTATCTACCATATCAAG 304
QY 353 CCACCTTAGCTTTAATAATTATGTTTCATTATTTC 389
Db 303 TTGGCTTCACTCCAGGCATTTTATTCAGTTTTTTAC 267

RESULT 9
US-10-027-632-49899/c
; Sequence 49899, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 49899
; LENGTH: 547
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-49899

Query Match      8.4%; Score 40.2; DB 13; Length 547;
Best Local Similarity 53.5%; Pred. No. 0.41;
Matches 84; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

QY 233 ACACAGAAAAGACACACAGCTCATCCACATGTAGACTGCTGTTATGCGTCGGC 292
Db 431 AAACAAAAAAGAAACCTCAGGCCAATATCCATGATGAACATTGACACAAAAATCCTCA 372
QY 293 TGAAGATAAGACTGACCCAGCCAGCAGCTAAAGAAAGAAATAATGCAAGTGTCTAGCT 352
Db 371 TTAATAATAAGTCAACCAATCCAGCAGCAATCAGAAAGTTTATCTACCATATCAAG 312
QY 353 CCACCTTAGCTTTAATAATTATGTTTCATTATTTC 389
```

```
US-10-027-632-72822/c
; Sequence 72822, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 72822
; LENGTH: 539
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-72822

Query Match      8.4%; Score 40.2; DB 16; Length 539;
Best Local Similarity 53.5%; Pred. No. 0.41;
Matches 84; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

QY 233 ACACAGAAAAGACACACAGCTCATCCACATGTAGACTGCTGTTATGCGTCGGC 292
Db 423 AAACAAAAAAGAAACCTCAGGCCAATATCCATGATGAACATTGACACAAAAATCCTCA 364
QY 293 TGAAGATAAGACTGACCCAGCCAGCAGCTAAAGAAAGAAATAATGCAAGTGTCTAGCT 352
Db 363 TTAATAATAAGTCAACCAATCCAGCAGCAATCAGAAAGTTTATCTACCATATCAAG 304
QY 353 CCACCTTAGCTTTAATAATTATGTTTCATTATTTC 389
Db 303 TTGGCTTCACTCCAGGCATTTTATTCAGTTTTTTAC 267

RESULT 8
US-10-027-632-312801/c
; Sequence 312801, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 312801
; LENGTH: 539
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-312801/c

Query Match      8.4%; Score 40.2; DB 16; Length 539;
Best Local Similarity 53.5%; Pred. No. 0.41;
Matches 84; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

QY 233 ACACAGAAAAGACACACAGCTCATCCACATGTAGACTGCTGTTATGCGTCGGC 292
Db 423 AAACAAAAAAGAAACCTCAGGCCAATATCCATGATGAACATTGACACAAAAATCCTCA 364
QY 293 TGAAGATAAGACTGACCCAGCCAGCAGCTAAAGAAAGAAATAATGCAAGTGTCTAGCT 352
Db 363 TTAATAATAAGTCAACCAATCCAGCAGCAATCAGAAAGTTTATCTACCATATCAAG 304
QY 353 CCACCTTAGCTTTAATAATTATGTTTCATTATTTC 389
Db 303 TTGGCTTCACTCCAGGCATTTTATTCAGTTTTTTAC 267

RESULT 9
US-10-027-632-49899/c
; Sequence 49899, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 49899
; LENGTH: 547
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-49899

Query Match      8.4%; Score 40.2; DB 13; Length 547;
Best Local Similarity 53.5%; Pred. No. 0.41;
Matches 84; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

QY 233 ACACAGAAAAGACACACAGCTCATCCACATGTAGACTGCTGTTATGCGTCGGC 292
Db 431 AAACAAAAAAGAAACCTCAGGCCAATATCCATGATGAACATTGACACAAAAATCCTCA 372
QY 293 TGAAGATAAGACTGACCCAGCCAGCAGCTAAAGAAAGAAATAATGCAAGTGTCTAGCT 352
Db 371 TTAATAATAAGTCAACCAATCCAGCAGCAATCAGAAAGTTTATCTACCATATCAAG 312
QY 353 CCACCTTAGCTTTAATAATTATGTTTCATTATTTC 389
```

Db 311 TTGGCTTCACTCCAGGCATTTTATTTCAAGTTTATTTTAC 275

## RESULT 10

US-10-027-632-49899/c  
; Sequence 49899, Application US/10027632  
; Publication No. US20030204075A9  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; Polymorphisms in the Human Genome  
; FILE REFERENCE: 108927.129  
; CURRENT APPLICATION NUMBER: US/10/027,632  
; PRIORITY FILING DATE: 2002-04-30  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 60/218,006  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 60/198,676  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: US 60/193,483  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: US 60/185,218  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/167,363  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: US 60/156,358  
; PRIOR FILING DATE: 1999-09-28  
; PRIOR APPLICATION NUMBER: US 60/146,002  
; PRIOR FILING DATE: 1999-08-09  
; NUMBER OF SEQ ID NOS: 325720  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 49899  
; LENGTH: 547  
; TYPE: DNA  
; ORGANISM: Human  
US-10-027-632-49899

Query Match 8.4%; Score 40.2; DB 16; Length 547;  
Best Local Similarity 53.5%; Pred. No. 0.41;  
Matches 84; Conservative 0; Mismatches 73; Indels 0; Gaps 0;  
QY 233 ACACGAGAAAGACACACAGCTCATCCACATGTAGATGTCGTTTGCCTCGGC 292  
Db 431 AACAAAAAAGAAACCTCAGGCCAATCCATGATGACATGACAAAAATCCTCA 372  
QY 293 TGAAGATAAGACTGACCCAGGCAGCAGCTATTAAGAGAAATATGCAAGTGTCTTAGCT 352  
Db 371 TTAATAATAGTCAACCAATCCAGCAGCAATCAGAAAGTTTATCTACCATATCAG 312  
QY 353 CCACCTTAGCTTTAATAATATGTTTCATTTATTTATTC 399  
Db 311 TTGGCTTCACTCCAGGCATTTTATTTCAAGTTTATTTTAC 275

## RESULT 11

US-10-318-819A-4  
; Sequence 4, Application US/10318819A  
; Publication No. US20040115645A1  
; GENERAL INFORMATION:  
; APPLICANT: C. Frank Bennett  
; APPLICANT: Kenneth W. Dobie  
; TITLE OF INVENTION: MODULATION OF DRK2 EXPRESSION  
; FILE REFERENCE: PFS-0069  
; CURRENT APPLICATION NUMBER: US/10/318,819A  
; PRIORITY FILING DATE: 2002-12-12  
; NUMBER OF SEQ ID NOS: 133  
; SEQ ID NO 4  
; LENGTH: 90798  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: 1349, 1350, 13454-13553, 28775-28874, 46851-46950, 52633-52732, 67991-  
; LOCATION: 68090, 85646-85745, 87254-87353, 89363-89462, 90660-90759

; OTHER INFORMATION: n = A, T, C or G  
US-10-318-819A-4

Query Match 7.8%; Score 37; DB 17; Length 90798;  
Best Local Similarity 51.5%; Pred. No. 60;  
Matches 85; Conservative 0; Mismatches 80; Indels 0; Gaps 0;  
QY 235 ACAGAAAAAGACACACAGCTCATCCACATGTAGATGTCGTTTGCCTCGGCTG 294  
Db 52993 AACAAAAAAGACCAATTTATTCCTAAAGAGTAAAGCATTTTCATGCTT 53052  
QY 295 AAGATAAGACTGACCCAGGCAGCAGCTAAAGAGAAATATGCAAGTGTCTTAGCTCC 354  
Db 53053 AAAATAGATGCTAAATCACCTTTTATAAAAAATTTACATTTCAAGTAATCCTAAGTAT 53112  
QY 355 ACTTAGCTTTAATAATATGTTTCATTTATTTATTTATTTATTTATTTATTTATTT 399  
Db 53113 AATTTTTTTTTTACTATCAGGAGTTAAACTACTCTCTCTTTT 53157

## RESULT 12

US-10-312-841-2/c  
; Sequence 2, Application US/10312841  
; Publication No. US20030186277A1  
; GENERAL INFORMATION:  
; APPLICANT: Epigenomics AG  
; TITLE OF INVENTION: Diagnose von bedeutenden genetischen Parametern innerhalb des MHC  
; FILE REFERENCE: E01/1208/NO  
; CURRENT APPLICATION NUMBER: US/10/312,841  
; CURRENT FILING DATE: 2002-12-30  
; NUMBER OF SEQ ID NOS: 2  
; SEQ ID NO 2  
; LENGTH: 3673778  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)  
; NAME/KEY: unsure  
; LOCATION: (379615)  
US-10-312-841-2

Query Match 7.8%; Score 37; DB 15; Length 3673778;  
Best Local Similarity 48.0%; Pred. No. 3.9e+02;  
Matches 106; Conservative 0; Mismatches 115; Indels 0; Gaps 0;  
QY 48 AATCAAGCAGAAAGAAATTAATACTTTTCAGATATGAAGCCGCTTTACAGTGGCCA 107  
Db 3196228 AATAAACCATAAATATAATATCTTAACACCATACCCTCCCATATAATTCATAACA 3196169  
QY 108 GCTAGCTATCACTGAAAGACAGCAGCAATGTTGTTCTCGATGACCCAGAACCAATCT 167  
Db 3196168 CCTATCAAAAAACGAAAAACATATAACAAAAAATAAATAATACGCCATAAAAAAACT 3196109  
QY 168 TTGCAGCAGATGTGAAGCAGCCAGGTGTCACAAAGCAGCACTCAGAAAAAGGATCTTC 227  
Db 3196108 CTTTTCCTTTTAAAAATCTTCGTATCTCTTAATAAAAAAATAAAAAACACAAATATC 3196049  
QY 228 TACCGACACAGAAAAAGACACACAGCTCATTCATCCAAACA 268  
Db 3196048 AATATAAATCAATATAAACTTCAATACATCTACACCCCAACA 3196008

## RESULT 13

US-10-221-714A-449/c  
; Sequence 449, Application US/10221714A  
; Publication No. US20040048254A1  
; GENERAL INFORMATION:  
; APPLICANT: OLEK, Alexander  
; APPLICANT: PIEFENBERG, Christian  
; APPLICANT: BERLIN, Kurt  
; TITLE OF INVENTION: Diagnosis of Diseases Associated with  
; tumor suppressor genes and oncogenes

FILE REFERENCE: 5013.1005  
CURRENT APPLICATION NUMBER: US/10/221,714A  
CURRENT FILING DATE: 2003-01-21  
PRIOR APPLICATION NUMBER: PCT/EP01/02955  
PRIOR FILING DATE: 2001-03-15  
PRIOR APPLICATION NUMBER: DE 10013847.0  
PRIOR FILING DATE: 2000-03-15  
PRIOR APPLICATION NUMBER: DE 10019058.8  
PRIOR FILING DATE: 2000-04-06  
PRIOR APPLICATION NUMBER: DE 10019173.8  
PRIOR FILING DATE: 2000-04-07  
PRIOR APPLICATION NUMBER: DE 10032529.7  
PRIOR FILING DATE: 2000-06-30  
PRIOR APPLICATION NUMBER: DE 10043826.1  
PRIOR FILING DATE: 2000-09-01  
NUMBER OF SEQ ID NOS: 540  
SEQ ID NO 449  
LENGTH: 6012  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)  
US-10-221-714A-449

Query Match 7.7%; Score 36.8; DB 13; Length 6012;  
Best Local Similarity 47.4%; Pred. No. 15;  
Matches 110; Conservative 0; Mismatches 122; Indels 0; Gaps 0;  
QY 34 AATTATTCAGATTAATCAAGGAGAGAAATTAATTAATCTTTTTCAGATATGAGCCCGC 93  
Db 861 ACTCAATTTTACAAAAAACAACAAACACTAAATACTTACCCCAAAATACAAAAACACC 802  
QY 94 TTTCAGTGGCCAGCTAGCTATCACTGAAAGACAGCAACATGGTGTCTCGATGCA 153  
Db 801 AAAACCACTCCCAAAACCTATACCCAAATACAAAAAATAATCTCAAAAA 742  
QY 154 CAGAACCACTCTTTGCGAGATGTGAAGACGCGAGTGTGTCACAGACGCACTCA 213  
Db 741 CAAAAAATATACGACAAAAAACAACAAAAAACCCTACATAAAAAAACCACCA 682  
QY 214 GAAAGGATCTCTTACCGACACAGAAAAACACACACACCTCATCA 265  
Db 681 ATAAAAACACACACACACACCCCAAAATTAACACACACACCCCAAAATCA 630

RESULT 14  
US-10-311-455-628/c  
Sequence 628, Application US/10311455  
Publication No. US20030143606A1  
GENERAL INFORMATION:  
APPLICANT: OLEK, Alexander  
APPLICANT: PIEPENBROCK, Christian  
TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determining Cytosine Methylation  
FILE OF INVENTION: cytosine methylation  
FILE REFERENCE: 5013.1014  
CURRENT APPLICATION NUMBER: US/10/311,455  
CURRENT FILING DATE: 2002-12-16  
PRIOR APPLICATION NUMBER: PCT/EP01/07537  
PRIOR FILING DATE: 2001-07-02  
PRIOR APPLICATION NUMBER: DE 10032529.7  
PRIOR FILING DATE: 2000-06-30  
PRIOR APPLICATION NUMBER: DE 10043826.1  
PRIOR FILING DATE: 2000-09-01  
NUMBER OF SEQ ID NOS: 2424  
SEQ ID NO 628  
LENGTH: 17211  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)  
US-10-311-455-628

Query Match 7.7%; Score 36.6; DB 15; Length 17211;  
Best Local Similarity 56.1%; Pred. No. 31;  
Matches 69; Conservative 0; Mismatches 54; Indels 0; Gaps 0;  
QY 17 AGATAATTCCTAATAGCAATTAATCAGATTAATCAAGGAGAGAAAGATTAATAACTCTTT 76  
Db 7211 AATAATCTCCCACTCTTCCTCGAATAATAATTAATAATTTTCAAAACTTT 7152  
QY 77 CAGAATATGAAGCCGCTTTTCAAGTGGCCAGCTAGCTATCACTGAAAGACAGCAAGAC 136  
Db 7151 TATATAATAATATCTCTATATAATTTCCCAACTCTCTATAACTATAAATTTCTCTATAA 7092  
QY 137 AAT 139  
Db 7091 AAT 7089

RESULT 15  
US-10-424-599-93424/c  
Sequence 93424, Application US/10424599  
Publication No. US20040031072A1  
GENERAL INFORMATION:  
APPLICANT: La Rosa Thomas J  
APPLICANT: Kovalic David K  
APPLICANT: Zhou Yihua  
APPLICANT: Cao Yongwei  
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with  
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
FILE REFERENCE: 38-21(53223)B  
CURRENT APPLICATION NUMBER: US/10/424,599  
CURRENT FILING DATE: 2003-04-28  
NUMBER OF SEQ ID NOS: 285684  
SEQ ID NO 93424  
LENGTH: 1141  
TYPE: DNA  
ORGANISM: Glycine max  
FEATURE:  
OTHER INFORMATION: Clone ID: PAT\_MRT3847\_55376C.1  
US-10-424-599-93424

Query Match 7.6%; Score 36.4; DB 13; Length 1141;  
Best Local Similarity 52.7%; Pred. No. 7.9;  
Matches 79; Conservative 0; Mismatches 71; Indels 0; Gaps 0;  
QY 313 GGCACGACATAAGAGAAATAATGCAAGTGTCTAGCTCCACTTTTACCTTTTATTAAT 372  
Db 1141 GCGCGCCCTTTTCAAGCAAAAAGACATTTATAACCAAGTCTGTCAAACTCACAAATTT 1082  
QY 373 ATGTTTCATTATTAATCTCTGCTTTTGTCTCTATATAAGAGCTTGTATTTCATTGA 432  
Db 1081 CAACTCCGGTTAACTTCATTTGCTTTTACGTATGAAGATTCATTTCTTTTGA 1022  
QY 433 AGCGAGAGCGCAACACACACAGAACCTC 462  
Db 1021 AAACCTAGATCAACAAAAAATATTTTC 992

Search completed: July 31, 2004, 09:31:32  
Job time : 340 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2004, 07:54:50 ; Search time 1998 Seconds  
(without alignments)  
7114.313 Million.cell updates/sec

Title: US-10-020-540A-1  
Perfect score: 476  
Sequence: 1 aactatgctgacgaagat.....aacctccctgcttacaacc 476

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST.\*

- 1: em\_estba.\*
- 2: em\_esthum.\*
- 3: em\_estlin.\*
- 4: em\_estnu.\*
- 5: em\_estov.\*
- 6: em\_estpl.\*
- 7: em\_estro.\*
- 8: em\_estc.\*
- 9: gb\_estl.\*
- 10: gb\_est2.\*
- 11: gb\_est3.\*
- 12: gb\_est4.\*
- 13: gb\_est5.\*
- 14: gb\_est6.\*
- 15: em\_estfun.\*
- 16: em\_estcom.\*
- 17: em\_gss\_hum.\*
- 18: em\_gss\_inv.\*
- 19: em\_gss\_pln.\*
- 20: em\_gss\_vrt.\*
- 21: em\_gss\_fun.\*
- 22: em\_gss\_man.\*
- 23: em\_gss\_mus.\*
- 24: em\_gss\_pro.\*
- 25: em\_gss\_rod.\*
- 26: em\_gss\_phg.\*
- 27: em\_gss\_vrl.\*
- 28: gb\_gsal.\*
- 29: gb\_gss2.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	DB ID	Description
C 1	54.8	11.5	994 13	EX414650
C 2	50.6	10.6	712 13	EX416727
C 3	44.4	9.3	886 29	CNS017XV
C 4	42.2	8.9	763 14	CA228227

5	41.4	8.7	1201	13	EX356232
C 6	40.8	8.6	919	29	CNS006S5
C 7	40.4	8.5	983	13	EX432706
C 8	40.4	8.5	1044	13	EX333633
C 9	40.2	8.4	648	13	EX417150
C 10	40.2	8.4	852	13	EX393687
C 11	40	8.4	1201	13	EX335650
C 12	39.8	8.4	775	29	CC484903
C 13	39.6	8.3	700	29	CE603042
C 14	39.6	8.3	1201	9	AL544517
C 15	39.2	8.2	885	13	EX425603
C 16	38.8	8.2	549	29	AG223026
C 17	38.8	8.2	870	28	AZ693178
C 18	38.8	8.2	1101	29	CNS00KX2
C 19	38.6	8.1	1101	29	CNS008SU
C 20	38.4	8.1	187	12	BG981764
C 21	38.2	8.0	682	29	CC820164
C 22	38.2	8.0	807	14	CF447957
C 23	38.2	8.0	834	29	CNS010ZS
C 24	38.2	8.0	958	10	BE777969
C 25	38	8.0	355	28	AQ044946
C 26	38	8.0	1059	13	EX359119
C 27	37.8	7.9	181	12	BI042556
C 28	37.6	7.9	826	29	CC572843
C 29	37.4	7.9	513	29	TA78C11P
C 30	37.4	7.9	593	29	CPA562119
C 31	37	7.8	527	29	CE497353
C 32	36.8	7.7	288	12	BJ392146
C 33	36.8	7.7	679	29	CE147815
C 34	36.8	7.7	688	28	B2452924
C 35	36.8	7.7	1201	13	EX461813
C 36	36.6	7.7	470	12	BU429271
C 37	36.6	7.7	477	12	BI324001
C 38	36.6	7.7	499	10	AW332217
C 39	36.6	7.7	511	10	AW335137
C 40	36.6	7.7	516	10	AW334251
C 41	36.6	7.7	522	10	AW335222
C 42	36.6	7.7	525	10	AW332905
C 43	36.6	7.7	534	10	AW334379
C 44	36.6	7.7	534	12	BU433770
C 45	36.6	7.7	550	10	AW332253

ALIGNMENTS

RESULT 1  
LOCUS BX414650/c  
DEFINITION BX414650 Homo sapiens THYMUS Homo sapiens cDNA clone CSOCAP001YN02  
3-PRIME mRNA sequence.  
ACCESSION BX414650  
VERSION BX414650.1 GI:30763455  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 994)  
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.  
TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished (2001)  
COMMENT Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: segre@genoscope.cns.fr, Web : www.genoscope.cns.fr  
Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 6015.f  
Contact : Feng Liang Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com/Invitrogen Corporation 1600 Faraday Avenue Genoscope sequence ID : CSOCAP001DG01NP1.  
FEATURES  
Location/Qualifiers  
1..994



from polyA+ mRNA using SuperScript Plasmid System Kit (Invitrogen). The double-strand cDNAs were fractionated in a sepharose CL-2B 40cm-columns and fragments sizing between 0.8 and 1.5 Kb were directionally cloned into the vector. Details of each source of RNA and library construction can be obtained at <a href="http://sucest.lad.ic.unicamp.br/public">http://sucest.lad.ic.unicamp.br/public</a>	
ORIGIN	
Query Match	8.9%; Score 42.2; DB 14; Length 763;
Best Local Similarity	54.2%; Presd. No. 21;
Matches 83;	Conservative 0; Mismatches 70; Indels 0; Gaps 0;
QY	317 ACCCTAAGAAGAAATAATGCAAGTGGTCTAGCTCCACTTAGCTTTAATAAATATGT 376
Db	695 AACATGAAGATGAAATGGGCGCAGCAAGGTGATGCTATCTTCACTAGACNATTAGAT 636
QY	377 TTCATTATATTCTCTGCTTTTGCTCTCTATATAAAGAGCTTGATTTTCATTTGAAGGC 436
Db	635 TTCCCAAGAAGTCTCCACACTTCTTAGCCACACAGAGTTGTTTTCATCTTCCTGTGAAGAG 576
QY	437 AGAGCGCAACACACACAGAACTCCCTGCTT 469
Db	575 ATATTGGCCCTAATAANAGAGTATCACCCCTTCTT 543
RESULT 5	
BOX56232	1201 bp mRNA linear EST 05-MAY-2000
LOCUS	BOX56232 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
DEFINITION	clone CS0D1009YF14 5-PRIME, mRNA sequence.
ACCESSION	BOX56232
KEYWORDS	BOX56232.1 GI:30370002
SOURCE	EST.
ORGANISM	Homo sapiens (human)
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS	Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE	Full-length cDNA libraries and normalization
JOURNAL	Unpublished (2001)
COMMENT	Contact: Genoscope Genoscope - Centre National de Sequencage BP 191 91006 EVRY cedex - France Email: <a href="mailto:segre@genoscope.cns.fr">segre@genoscope.cns.fr</a> , Web : <a href="http://www.genoscope.cns.fr">www.genoscope.cns.fr</a> Library was constructed by life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 6268.r For more information about this cluster, see <a href="http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0D1009DC07QP1&amp;cluster=6268.r">http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0D1009DC07QP1&amp;cluster=6268.r</a> Feng liang Email : <a href="mailto:fliang@lifetech.com">fliang@lifetech.com</a> URL : <a href="http://fulllength.invitrogen.com/InvitrogenCorporation1600ParadiseAvenueGenoscopeSequenceID:CS0D1009DC07QP1">http://fulllength.invitrogen.com/InvitrogenCorporation1600ParadiseAvenueGenoscopeSequenceID:CS0D1009DC07QP1</a> Location/Qualifiers 1..1201 /organism="Homo sapiens" /mol_type="mRNA" /db_xref="taxon:9606" /clone="CS0D1009YF14" /tissue="PLACENTA" /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED" /note="First strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."
FEATURES	
source	



```
QY 401 TCTCTATATAGAGCTGTGTTTTCATTTTGAAGCAGAGCGGNACAC 450
DB 487 TWTHHHHHHMCVTTTTTTTTTTTTTTTTTTTGGACAAGCTTAAACAC 438

RESULT 8
BX333633/c
LOCUS
DEFINITION BX333633 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens
CDNA clone CS0DC028Y011 3-PRIME, mRNA sequence.
ACCESSION BX333633
VERSION
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)

REFERENCE
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1. (bases 1 to 1044)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 10303.r for
more information about this cluster, see
http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0DC028AH06NP1&cluster=10303.r. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ InvitroGen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DC028AH06NP1.

FEATURES
source
1..1044
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DC028Y011"
/tissue_type="NEUROBLASTOMA COT 25-NORMALIZED"
/clone_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR.V
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
Query Match 8.5%; Score 40.4; DB 13; Length 1044;
Best Local Similarity 30.3%; Pred. No. 51;
Matches 72; Conservative 57; Mismatches 109; Indels 0; Gaps 0;

QY 5 ATGCTGATGACAGATATTTCTAAGCAATTATTGAGAAATTAATCAAGGAGAAAGAT 64
DB 591 ATTHGSAASAASACTTCTCCCAAGSKSAKSAKSAKSAKSAKSAKSAKSAKSAK 532
QY 65 TAAATACTTTTCAGATATGACCGCGTTTACAGTGGCCAGCTAGCTATCACTGAA 124
DB 531 ACAMAAAKKAKKASSASATAAATACTCTAMWKCCTTCSAASAKMSAGSKMKVASAM 472
QY 125 AGACAGCAAGCAATGTGTCTGATGACAGCAACCAATCTTTGACGAGATGTGAAG 184
DB 471 TAAAGSGAGGGAAMAKMTATCKGTAASSAMAAATTTAATTASCMAKMSAKATWKS 412
QY 185 CAGCAGAGTGGTCCACAGAGCGCACTCAGAAAGGCACTTTTACCGACACAGAAA 242
DB 411 AAMSSCVGGGTWGMMAACSGSGSGCCGAMAGCMACMWASSASCAMSGGAGAAA 354

RESULT 9
BX417150/c
LOCUS
DEFINITION BX417150 Homo sapiens PLACENTA Homo sapiens cDNA clone CS0DE005YMC1
5-PRIME, mRNA sequence.

ACCESSION BX417150
VERSION
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
```

```
ACCESSION BX417150
VERSION BX417150.1 GI:30658341
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1. (bases 1 to 648)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ InvitroGen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DE005AG01QPI.

FEATURES
source
1..648
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DE005YMC1"
/tissue_type="PLACENTA"
/clone_lib="Homo sapiens PLACENTA"
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."

ORIGIN
Query Match 8.4%; Score 40.2; DB 13; Length 648;
Best Local Similarity 23.9%; Pred. No. 60;
Matches 74; Conservative 72; Mismatches 163; Indels 0; Gaps 0;

QY 122 AAAAGCAGCAGACAAATGGTGTCTCGATGACACAGAACCACTTTTCGACGATGTG 181
DB 588 AAAAGAAAAAANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 529
QY 182 AAGCAGCAGAGTGTGTCACAGACGCACTCAGAAAAGGCACTTCTCCGACACAGAAA 241
DB 528 MMTMMMTMMMTMMMTMMMTMMMTMMMTMMMTMMMTMMMTMMMTMMMTMMMTMM 469
QY 242 AAGAACACACAGCTCATCATCCACATGACACTGCTGTTATGCGTGAAGATAA 301
DB 468 MMTMMMTMMMTMMMTMMMTMMMTMMMTMMMTMMMTMMMTMMMTMMMTMMMTMM 409
QY 302 GACTGACCCCGCCAGCAGCACTAAAGAGAAATAATGCAAGTGTCTCTAGCTCCACTTTAG 361
DB 408 MMTMAAAAAMVWVCCCCCCCCAGGGRAAAAMAAAAMVWVCCCCCCCCCMHMTWTTT 349
QY 362 CTTTAATAATGTTTTCATTTATTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 421
DB 348 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 289
QY 422 TTTTCATTT 430
DB 288 HTTYTTTTT 280

RESULT 10
BX393687/c
LOCUS
DEFINITION BX393687 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens
CDNA clone CS0DC001YG07 5-PRIME, mRNA sequence.
ACCESSION BX393687
VERSION
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
```



Clones are derived from the bovine BAC library CHORI-240 (<http://www.chori.org/bacpac/bovine240.htm>). For BAC library availability, please contact Pieter de Jong ([pdejong@mail.choi.org](mailto:pdejong@mail.choi.org)). Clones may be purchased from BACPAC Resources (<http://www.chori.org/bacpac/ordering.information.htm>). This work was undertaken as part of the International Bovine BAC Mapping Consortium (IBBMC) by CSIRO Livestock Industries, Australia and the British Columbia Genome Sciences Centre, Canada.

Plate: 314 row: I column: 5  
Seq primer: SP6  
Class: BAC ends.

#### FEATURES

Location/Qualifiers  
1..775

/organism="Bos taurus"  
/mol\_type="genomic DNA"  
/strain="Breed: Hereford"  
/db\_xref="taxon:9913"  
/clone="CH240\_31415"  
/sex="Male"  
/cell\_type="Blood"  
/clone\_lib="CHORI-240"  
/notes="Vector: pTARBAC1.3; Site 1: MboI; Site 2: MboI; Hereford bull Li Domino 9375; CHORI-240 Bovine BAC library (Male) produced by Pieter de Jong"

#### ORIGIN

Query Match 8.4%; Score 39.8; DB 29; Length 775;  
Best Local Similarity 60.7%; Pred. No. 73;  
Matches 65; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

QY 333 TAATGCAAGTGGTCCAGTCCACTTAGCTTTAATAATATGTTTCATATATCTCT 392  
Db 385 TCATACAGACATCTAGTTCATATATATATATATATATATATATATATATAT 326  
QY 393 GCATTTCCTCTATATAAGAGCTGTATTTTCATTGAAGCACA 439  
Db 325 GCATTTCATGTGTAGAAATCTCTCTTGTGGAGTCATA 279

#### RESULT 13

CE603042/c  
LOCUS  
DEFINITION  
tigr-gss-dog-17000356692102 Dog Library Canis familiaris genomic,  
genomic survey sequence.

ACCESSION  
CE603042  
VERSION  
CS603042.1 GI:36919881  
KEYWORDS  
GSS.  
SOURCE  
Canis familiaris (dog)  
ORGANISM  
Canis familiaris

REFERENCE  
1 (bases 1 to 700)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
Rusch, D.B., Deicher, A.L., Pop, M., Wang, W., Fraser, C.M. and  
Venter, J.C.

TITLE  
The dog genome: survey sequencing and comparative analysis

JOURNAL  
Science 301 (5641), 1998-1903 (2003)

MEDLINE  
22875432

PUBMED  
14512627

COMMENT  
Contact: Kirkness EF

The Institute for Genomic Research  
Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,  
Rockville, MD 20850, USA  
Tel: 301-838-0200  
Fax: 301-838-0208  
Email: [ekirknes@tigr.org](mailto:ekirknes@tigr.org)  
Class: shotgun.

Location/Qualifiers  
1..700

/organism="Canis familiaris"  
/mol\_type="genomic DNA"  
/strain="Standard Poodle"  
/db\_xref="taxon:9615"

#### FEATURES

source

/clone\_lib="Dog Library"  
/note="Site 1: BstXI; Libraries were prepared from  
peripheral blood"

#### ORIGIN

Query Match 8.3%; Score 39.6; DB 29; Length 700;  
Best Local Similarity 58.5%; Pred. No. 82;  
Matches 69; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

QY 351 CTCACATTTAGCTTTAATAATATGTTTCATATATATCTCTGTTTGCCTCTATATA 410  
Db 644 CTTTAATTTGGGTTTATCTAATGTTCTTATGATTAGATGTTGTTCTGCAATTTCTATA 585  
QY 411 AAGAGCTTGTATTTTCATTTCGAGGAGGCGGAACACACACAGACACCTCCCTGCT 468  
Db 584 AATAATGTGTGTTTCTTATGATGCATTTAAATCAGACATAGAGGCACCTGGGT 527

#### RESULT 14

AL544517/c

LOCUS

DEFINITION

clone CS01021Y122 5-PRIME, mRNA sequence.

ACCESSION

AL544517

VERSION

AL544517.2 GI:31266360

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

Li, W.B., Gruber, C., Jessee, J. and Polayes, D.

Full-length cDNA libraries and normalization

Unpublished (2001)

COMMENT

On Feb 15, 2001 this sequence version replaced gi:12876997.

Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: [segref@genoscope.cns.fr](mailto:segref@genoscope.cns.fr), Web : [www.genoscope.cns.fr](http://www.genoscope.cns.fr)

Library was constructed by Life Technologies, a division of

Invitrogen. This sequence belongs to sequence cluster 3154.f For

more information about this cluster, see

[http://www.genoscope.cns.fr/](http://www.genoscope.cns.fr/cgi-bin/cluster.cgi?seq=CS0D1021BE11Q1&cluster=3154.f)

Feng Liang Email : [fliang@lifetech.com](mailto:fliang@lifetech.com) URL :

<http://fulllength.invitrogen.com/> Invitrogen Corporation 1600

Faraday Avenue Genoscope sequence ID : CS0D1021BE11Q1.

Location/Qualifiers

1..1201

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="CS0D1021Y122"

/tissue type="PLACENTA"

/clone\_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"

/note="1st strand cDNA was primed with a NotI-oligo (dT)

primer. Five prime end enriched, double-strand cDNA was

digested with NotI and cloned into the NotI and EcoRV

sites of the pCMVSPORT 6 vector. Library was normalized."

#### ORIGIN

Query Match 8.3%; Score 39.6; DB 9; Length 1201;  
Best Local Similarity 31.9%; Pred. No. 76;  
Matches 102; Conservative 58; Mismatches 160; Indels 0; Gaps 0;

QY 122 AAAGACAGCAAGACATGGTCTTCGATGCACAGAACACATCTTTGACGAGATGTG 181  
Db 1117 MWMWMMWMMWMAAAANDWMMWYCCCGGGSSNAAMCCCGGGKKKKKYYCAAAMWY 1058  
QY 182 AAGCAGCCAGATGTCACAGAGCGACTCAGAAAAGGCACTTCTACCCACACAGAA 241  
Db 1057 CCCCMAAAARRRGGKCCYMAAAAMCMAAAARATTTTCCCHMMMRDRTWAA 998





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2004, 08:52:25 ; Search time 2305 Seconds  
(without alignments)  
8950.666 Million cell updates/sec

Title: US-10-020-540A-1  
Perfect score: 476  
Sequence: 1 aactatgctgatgacaagat.....aacctccctgcttacaacc 476

Scoring table:  
OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 3470272 seqs, 21671516995 residues

Word size : 0

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

GenEmbl.\*

1: gb.ba.\*  
2: gb.htg.\*  
3: gb.in.\*  
4: gb.om.\*  
5: gb.ov.\*  
6: gb.pat.\*  
7: gb.ph.\*  
8: gb.pl.\*  
9: gb.pr.\*  
10: gb.ro.\*  
11: gb.sts.\*  
12: gb.sy.\*  
13: gb.un.\*  
14: gb.vi.\*  
15: em.ba.\*  
16: em.fun.\*  
17: em.hum.\*  
18: em.in.\*  
19: em.mu.\*  
20: em.om.\*  
21: em.ov.\*  
22: em.or.\*  
23: em.pat.\*  
24: em.ph.\*  
25: em.pl.\*  
26: em.ro.\*  
27: em.sts.\*  
28: em.un.\*  
29: em.vi.\*  
30: em.htg\_hum.\*  
31: em.htg\_inv.\*  
32: em.htg\_other.\*  
33: em.htg\_mus.\*  
34: em.htg\_pln.\*  
35: em.htg\_rod.\*  
36: em.htg\_man.\*  
37: em.htg\_vrt.\*  
38: em.sy.\*  
39: em.htgo\_hum.\*  
40: em.htgo\_mus.\*  
41: em.htgo\_other.\*

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	476	100.0	476	6	AX474089	AX474089 Sequence
2	476	100.0	518	6	AX474090	AX474090 Sequence
3	476	100.0	7876	14	SVBVCOMGN	X97304 Strawberry
4	435	91.4	950	14	AF331666	AF331666 Strawberry
5	369	77.5	651	6	AX474091	AX474091 Sequence
6	369	77.5	701	6	AX474092	AX474092 Sequence
c 7	21	4.4	41339	9	AC005490	AC005490 Homo sapi
8	21	4.4	192672	2	AC121285	AC121285 Mus muscu
c 9	21	4.4	221190	2	BX571704	BX571704 Danio rer
10	20	4.2	1046	8	AK061232	AK061232 Oryza sat
11	20	4.2	1490	8	SCYKLO41W	Z28041 S.cerevisia
12	20	4.2	1557	6	AX008583	AX008583 Sequence
13	20	4.2	1804	8	SCYKLO40C	Z28040 S.cerevisia
14	20	4.2	2204	8	AF307842	AF307842 Chlamydom
c 15	20	4.2	2709	10	BC051144	BC051144 Mus muscu
c 16	20	4.2	2924	10	BC021479	BC021479 Mus muscu
c 17	20	4.2	6576	8	YSCSEC7	J03918 Yeast (S.ce
18	20	4.2	9539	8	AF307843	AF307843 Chlamydom
19	20	4.2	13631	8	SC9489	Z47813 S.cerevisia
20	20	4.2	17564	8	SCELMIAA	X71621 S.cerevisia
21	20	4.2	75594	9	AF003093	AF003093 Homo sapi
c 22	20	4.2	83007	8	AC123575	AC123575 Medicago
23	20	4.2	86314	9	AC109592	AC109592 Homo sapi
24	20	4.2	106273	9	HSJ1018A4	AL109612 Human DNA
25	20	4.2	110000	2	AC111399.2	Continuation (3 of
c 26	20	4.2	141529	2	AC025330	AC025330 Homo sapi
c 27	20	4.2	151494	2	AC132889	AC132889 Rattus no
28	20	4.2	156945	10	AC105944	AC105944 Mus muscu
c 29	20	4.2	159857	14	AF170722	AF170722 Rabbit fl
c 30	20	4.2	160847	8	AC120307	AC120307 Oryza sat
c 31	20	4.2	163120	9	AC011454	AC011454 Homo sapi
32	20	4.2	171493	2	AC067933	AC067933 Homo sapi
c 33	20	4.2	173967	9	AC073343	AC073343 Homo sapi
c 34	20	4.2	174526	2	AC101807	AC101807 Mus muscu
35	20	4.2	176026	9	AC092661	AC092661 Homo sapi
c 36	20	4.2	184506	10	AC112701	AC112701 Mus muscu
37	20	4.2	188412	2	AC119098	AC119098 Rattus no
c 38	20	4.2	192843	10	AL607124	AL607124 Mouse DNA
c 39	20	4.2	197352	2	AC103365	AC103365 Mus muscu
40	20	4.2	205621	10	AC079082	AC079082 Mus muscu
41	20	4.2	207922	10	AL450399	AL450399 Mouse DNA
42	20	4.2	213875	10	AC096975	AC096975 Rattus no
c 43	20	4.2	214831	2	AC125984	AC125984 Rattus no
c 44	20	4.2	220446	2	AC129443	AC129443 Rattus no
45	20	4.2	223979	2	AC096151	AC096151 Rattus no

ALIGNMENTS

RESULT 1	AX474089	476 bp	DNA	linear	PAT 09-AUG-2002
LOCUS	Sequence 1 from Patent WO0240691.				
DEFINITION	AX474089				
ACCESSION	AX474089.1	GI:22208238			
VERSION	Strawberry vein banding virus (SVBV)				
KEYWORDS	Strawberry vein banding virus				
SOURCE	Viruses; Retroid viruses; Caulimoviridae; Caulimovirus.				
ORGANISM	1				
REFERENCE	Wu, G. and McBride, K.				
AUTHORS	Strawberry vein banding virus (svbv) promoter				
TITLE	Patent: WO 0240691-A 1 23-MAY-2002;				
JOURNAL	Maxygen, Inc. (US)				

Pred. No. is the number of results predicted by chance to have a

```

FEATURES             Location/Qualifiers
     source
1..476
/organism="Strawberry vein banding virus"
/mol_type="unassigned DNA"
/strain="E3"
/db_xref="taxon:47903"

ORIGIN
Query Match      100.0%; Score 476; DB 6; Length 476;
Best Local Similarity 100.0%; Pred. No. 1.4e-261;
Matches 476; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AACTATGCTGATGACAAGATAATTTCTTAATAGCAATTAATTCAGAATTAATCAAGAGAAA 60
DB 1 AACTATGCTGATGACAAGATAATTTCTTAATAGCAATTAATTCAGAATTAATCAAGAGAAA 60
QY 61 GAATTAATAACTCTTTTCAAGATATGAAGCCCGCTTTTACAAGTGGCCAGCTAGCTATCACT 120
DB 61 GAATTAATAACTCTTTTCAAGATATGAAGCCCGCTTTTACAAGTGGCCAGCTAGCTATCACT 120
QY 121 GAAAGACAGCAGCAATGCTGCTCGATGACCAAGAACCAACATCTTTGACGAGATGT 180
DB 121 GAAAGACAGCAGCAATGCTGCTCGATGACCAAGAACCAACATCTTTGACGAGATGT 180
QY 181 GAAGCAGCCAGAGTGGTCCACAAGACGCACTCAGAAAAGGCAATCTTACCGACACAGAA 240
DB 181 GAAGCAGCCAGAGTGGTCCACAAGACGCACTCAGAAAAGGCAATCTTACCGACACAGAA 240
QY 241 AAAGCAACACCAAGCTCATCTCAACATGTAGACTGTGCTTATCGTGGCTGAAGATA 300
DB 241 AAAGCAACACCAAGCTCATCTCAACATGTAGACTGTGCTTATCGTGGCTGAAGATA 300
QY 301 AGACTGACCCAGGCGGACACACACACACACACACACACACACACACACACACACACAC 360
DB 301 AGACTGACCCAGGCGGACACACACACACACACACACACACACACACACACACACACAC 360
QY 361 GCTTTAATAATTAATTTTCAATTAATTTCTCTGCTTTTCTCTCTATATAAAGAGCTTGT 420
DB 361 GCTTTAATAATTAATTTTCAATTAATTTCTCTGCTTTTCTCTCTATATAAAGAGCTTGT 420
QY 421 ATTTTCATTTGAAGCGGACGACACACACACACACACACACACACACACACACACACAC 476
DB 421 ATTTTCATTTGAAGCGGACGACACACACACACACACACACACACACACACACACACAC 476

RESULT 2
AX474090
LOCUS             518 bp      DNA      linear      PAT 09-AUG-2002
DEFINITION       Sequence 2 from Patent WO0240691.
ACCESSION        AX474090
VERSION          AX474090.1
KEYWORDS         Strawberry vein banding virus (SVBV)
SOURCE           Strawberry vein banding virus
ORGANISM         Viruses; Retroid viruses; Caulimoviridae; Caulimovirus.
REFERENCE        1 Wu,G. and McBride,K.
AUTHORS          Patent: WO 0240691-A 2 23-MAY-2002;
JOURNAL          Maxygen, Inc. (US)
FEATURES             Location/Qualifiers
     source
1..518
/organism="Strawberry vein banding virus"
/mol_type="unassigned DNA"
/strain="E3"
/db_xref="taxon:47903"

ORIGIN
Query Match      100.0%; Score 476; DB 6; Length 518;
Best Local Similarity 100.0%; Pred. No. 1.4e-261;
Matches 476; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AACTATGCTGATGACAAGATAATTTCTTAATAGCAATTAATTCAGAATTAATCAAGAGAAA 60

FEATURES             Location/Qualifiers
     source
1..7876
/organism="Strawberry vein banding virus"
/mol_type="genomic DNA"
/species="host="Pragaria sp."
/db_xref="taxon:47903"
/cloned="pSVBV-E3"
/gene="ORF I"
67..1053
/gene="ORF I"
67..1053
/codon_start=1
/protein_id="CAA65972.1"
/db_xref="GI:1360609"
/db_xref="GOA:Q88438"
/db_xref="SPTREMBL:Q88438"
/translation="MSEBEIRMDQGGHDBYIFEEGTYAHDVAIDSTLLKEIEKKD

SVBVCOMGN
Strawberry vein banding virus complete genome.
X97304
X97304.1 GI:1360608
complete genome; ORF1; ORFII; ORFIII; orfIV; orfV; ORFVI.
Strawberry vein banding virus (SVBV)
SOURCE           Strawberry vein banding virus
ORGANISM         Viruses; Retroid viruses; Caulimoviridae; Caulimovirus.
REFERENCE        1 Petrzik,K., Benes,V., Mraz,I., Honetslegrova-Franova,J., Ansoerge,W.
AUTHORS          and Spak,J.
TITLE            Strawberry vein banding virus--definitive member of the genus
JOURNAL          Caulimovirus
MEDLINE          Virus Genes 16 (3), 303-305 (1998)
PUBMED           98318753
REFERENCE        2 (bases 1 to 7876)
AUTHORS          Petrzik,K.
TITLE            Direct Submission
JOURNAL          Submitted (03-APR-1996) K. Petrzik, Institute of Plant Molecular
Biology, Dept. Plant Virology, Braniscovska 31, Ceske Budejovice,
CZ-370 05, Czech Republic
FEATURES             Location/Qualifiers
     source
1..7876
/organism="Strawberry vein banding virus"
/mol_type="genomic DNA"
/species="host="Pragaria sp."
/db_xref="taxon:47903"
/cloned="pSVBV-E3"
/gene="ORF I"
67..1053
/gene="ORF I"
67..1053
/codon_start=1
/protein_id="CAA65972.1"
/db_xref="GI:1360609"
/db_xref="GOA:Q88438"
/db_xref="SPTREMBL:Q88438"
/translation="MSEBEIRMDQGGHDBYIFEEGTYAHDVAIDSTLLKEIEKKD

```

```
LELTSBEVFKTSLMKFLKARKNICISCVSREYPIEIAQANGLEIPIFFNREBIES
KXVLKPERKIDFTHIGSVKIMIKSTRTIGDPSISVALLDRMKNAKDAVFGVK
GNLSYKLIFTYKPKISVLSRPTINKTILTAHFPEKEBELMEHGNHPYITISYKIVTL
SNHSHLEPRKPEPICIDDLFTSVGKISQAPLOEITPIENYWRMSLNGTSRELLGERP
RMVIEDDEEIPLDRIETALARASOSRMGLRPLUEDLRTTSKINILAHQL"
1058..1543
/gene="ORF II"
1058..1543
/gene="ORF II"
/codon_start=1
/protein_id="CAA65973.1"
/db_xref="GI:1360610"
/translation="MSFRPEPHIYKKQYLTNTQNVERGEPEYLYVVGRIEGCLKH
LNNIVCGNIHMNYIYIKTIGIKTSIYRNPEDNSPGLSLILGKPOESSFONPLE
KKIDNLSKIDRLUGSNTTGTETEKLAIDLNSKISELERKISSLTDDIMNLLKLDLDDK
C"
1543..1896
/gene="ORF III"
1543..1896
/gene="ORF III"
/codon_start=1
/protein_id="CAA65974.1"
/db_xref="GI:1360611"
/db_xref="SPTREMBL:Q88440"
/translation="MGSVTKTERIYDEMLKMEERSRLERQEQDHFSEKIDQASAIL
LABIRKIQSKLETCDCKEILDAKQDRTKKNGSPSLNLDALDLAHKDNSPKRNPIR
GTENWHPQDLKLTW"
1890..3314
/gene="ORF IV"
1890..3314
/gene="ORF IV"
/codon_start=1
/protein_id="CAA65969.1"
/db_xref="GI:1360612"
/db_xref="GOA:Q88441"
/db_xref="SPTREMBL:Q88441"
/translation="MYSRREPLFOLEDEHPMDIIOVLSLHDLDDCOEKLQVL
VAKEMIDYSSDEDETRNVKVIKEDEETVRPNKRKGNSSSQNYRYDIPBEYI
PNRKTGLNTSLDLDCASNRROEILBENDNEMKLIITBKALJNTDFDLITLAKKTV
GNAKGLTSLHTFAFQAOSTGEEFTLTNSFTYFVGTNYLTQGTREKEXAGEAR
NRVLQICNLCSLFFCDYETNLKPIEEMPKYIEBYIRKIPVGMVELEVESKQ
DSTTKGSLYAHNLIKAYVEKCKSLIKKEIRNMCCPKFSPETQYCGCKTISHKA
KKQKYQYKYYKYLKPKRWNSRRKYSRKLFRKRDQKSETSOQSPBEKKKFCPO
KTCRCWICNEIGHCAKDCRNKSANHNKILIELOSLQLEPVPDINELKIEKFWEK
EVSESESEISESDESSEDLLE"
3406..5532
/gene="ORF V"
3406..5532
/gene="ORF V"
/codon_start=1
/protein_id="CAA65970.1"
/db_xref="GI:1360613"
/db_xref="GOA:Q88442"
/db_xref="SPTREMBL:Q88442"
/translation="MENNLQLELHLEBGEWMLTDEDFKLOQNEDEBQLLESVDSE
ERPFLTNTPNSTYIRGNFYFKYKSLDVLVDTCASCTANKHYIPEEFWNAKN
PIRARIANDIMFNKVAELMOVLADETILPLACATTKGDITIGNNFCRLYEPFV
QYKDMITFDHGDRAVSKYTKAYFHLPGFLESKKVGSSTSPNENITPTVINGDN
ISKIFGEEIESEQLFSTISAYTEVEKLDDISCSBHPDLSRINKGRFQAALLDTN
KIINKCPMQSPQDREBFTQIIEELKGIIRPSKSPHSPAFMVRNHAIRKGRARM
VINYKLNDRTKGIDYLLPNKQQLORIGKTFYSSFDCKSGFQVRLAPETIQLTAF
SCPOGHYELMWLPFGLLQPAIFORHMDSELSNMPQFCAYVYDDIIVFSKTEEEHLG
HVKIVLNRCKALGVLKSKAOLCKTTFNPLGIVIBERNLKVQSHGLHLHAPPDOLS
DRNALORPLGLNVIISAYFPKIANLSPLOVKLKEITWSWTEKDTETVRKIKSLVKT
LPDNLSPDEKPIEFCDASDDHWGAILKALPEGKEVICRYASGTFKPAEKNYHNE
KEILSILIKAFIAYILPYKFLVTRTNAAIFVRNINAGDIQAQSLVRWQNALREY
SFDIEHSVGGQNVLIADIMTRELAKT"
5544..7088
/gene="ORF VI"
5544..7088
/gene="ORF VI"
/codon_start=1
```

```
/protein_id="CAA65971.1"
/db_xref="GI:1360614"
/db_xref="GOA:Q88443"
/db_xref="SPTREMBL:Q88443"
/translation="MEQVALSYESLHQNNEVKLSIYDLEDQISAKRYBEDLLERKL
RRHSAVKSQODNITGILPREAVIIPPRKLEVTQQLGEGGTQSTELSSKQPEEGKP
NNSQSYTGIRPQFQSDNQTTKVSAPQVELSPQAPLSGSKTYIVYDGNQGIYDS
WALVAPIKQNYHHQFTRKEAELALQTYVQHKRPIEQHQPKQAQFPESDYKK
ALASTKAAPKNLVGLRMFKQPKITEKSKVPERPLVTQDHFTHNVLVQTHDNCEE
HKFTVDSKTLISLYGFTIEGADPEVVRBAFHCGLISLIPGENLKEISILLPKDLYRAIK
NYRTEVCSAOSRPYILRITSFLEWENDLILIPVHTEVGLSNGKIPDKSKYKKTATV
EYLFETALRPRKRIKAVYNQCTKIFADSKKLNLYADDKILLISNYSLLIEKELITLIS
EYEFRTSGQALITEKTAQRQCLDPAPEPLHCSRCEAAARVHVHKTSEKASSTDTTEKDNH
SSSSNM"
ORIGIN
Query Match 100.0% Score 476; DB 14; Length 7876;
Best Local Similarity 100.0%; Pred. No. 1.5e-261;
Matches 476; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 AACTATGCTGATGACAGATAATTTCTAATAAGCAATTAATCAGAAATTAATCAAGAGAAA 60
DB 6816 AACTATGCTGATGACAGATAATTTCTAATAAGCAATTAATCAGAAATTAATCAAGAGAAA 6875
QY 61 GAATTAATAACTCTTTCAGRATATGAGCCCGCTTTCAAGTGGCCAGCTAGCTATCACT 120
DB 6876 GAATTAATAACTCTTTCAGRATATGAGCCCGCTTTCAAGTGGCCAGCTAGCTATCACT 6935
QY 121 GAAAAGACAGCAAGACAATGGTGTCTCGATGACACAGAACACACATCTTTTGACGAGATGT 180
DB 6936 GAAAAGACAGCAAGACAATGGTGTCTCGATGACACAGAACACACATCTTTTGACGAGATGT 6995
QY 181 GAAGCAGCAGAGTGTCTCAAGAGCGACTCAGAAAAGGCACTTCTTACGACACAGAA 240
DB 6996 GAAGCAGCAGAGTGTCTCAAGAGCGACTCAGAAAAGGCACTTCTTACGACACAGAA 7055
QY 241 AAAGCAACACACAGCTCATCATCAACATGTAGACTGCTTATGCGTCGGCTGAAGATA 300
DB 7056 AAAGCAACACACAGCTCATCATCAACATGTAGACTGCTTATGCGTCGGCTGAAGATA 7115
QY 301 AGACTGACCCAGGCGACACTAAGAGAGAAATAATCAAGTGGTCTTAGCTCCACTTTA 360
DB 7116 AGACTGACCCAGGCGACACTAAGAGAGAAATAATCAAGTGGTCTTAGCTCCACTTTA 7175
QY 361 GCTTTTAATAATATGCTTTTCAATTAATTTCTGCTTTTGTCTCTCTATATAAGAGCTTGT 420
DB 7176 GCTTTTAATAATATGCTTTTCAATTAATTTCTGCTTTTGTCTCTCTATATAAGAGCTTGT 7235
QY 421 ATTTTCATTTGAAGCGAGCGGCAACACACACAGAACCTCCCTGCTTACAAACC 476
DB 7236 ATTTTCATTTGAAGCGAGCGGCAACACACACAGAACCTCCCTGCTTACAAACC 7291
RESULT 4
AF331666 950 bp DNA linear VRL 15-MAR-2001
LOCUS Strawberry vein banding virus ORF VI (ORF VI) gene, partial cds and
DEFINITION viral promoter sequence.
ACCESSION AF331666
VERSION AF331666.1 GI:13345788
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 (bases 1 to 950)
AUTHORS Wang, Y., Gaba, V., Wolf, D., Xia, X.D., Zelcer, A. and Gal-On, A.
TITLE Identification of a novel plant virus promoter using a potyvirus
JOURNAL Infectious clone
MEDLINE Virus Genes 20 (1), 11-17 (2000)
PUBMED 2027369
REFERENCE 10766302
AUTHORS Wang, Y., Gaba, V., Wolf, D., Xia, X.D., Zelcer, A. and Gal-On, A.
```

TITLE Direct Submission  
JOURNAL Submitted (21-DEC-2000) Plant Virology Dept., The Volcani  
Center-ARO, Bet Dagan 50-250, Israel  
FEATURES  
Location/Qualifiers  
1..950  
/organism="Strawberry vein banding virus"  
/proviral  
/mol\_type="genomic DNA"  
/db\_xref="taxon:47903"  
c1..788  
/gene="ORF VI"  
c1..788  
/gene="ORF VI"  
/codon\_start=3  
/product="ORF VI"  
/protein\_id="AAK19514.1"  
/db\_xref="GI:13345789"  
/translation="PLVTKDHFHNYLVAQTHDNCBEHKFTVDSKTLISYGFIEGAD  
PEYREAFCGLISLIYFENLKEISLIPKOLYRAIKNTRVCSAGSRPIYLRTST  
PLEWENDILIPYHWEIVGLSNKIPMDKSYKKTATVEYLPETALRPKIKAVYNO  
TKIFADSKKILNYADDKILISNYSSELTKEKELITLSEYEAFTSGQLAITEKTARQW  
CLDAPEPHLCRCEAARVVHKTSHSEKASSTDEKDNHSSSNM"  
623..950  
promoter  
/note="sufficient to act as a viral promoter"  
ORIGIN  
Query Match 91.4%; Score 435; DB 14; Length 950;  
Best Local Similarity 100.0%; Pred. No. 5.1e-238;  
Matches 435; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 AACTATGCTGATGACAGCAATATTTCTAATAGCAATATTTCAGAAATTAATCAAGAGAAA 60  
Db 516 AACTATGCTGATGACAGCAATATTTCTAATAGCAATATTTCAGAAATTAATCAAGAGAAA 575  
QY 61 GAATTAATAACCTTTTCAGAAATATCAAGCCGCTTTTACAAGTGGCCAGCTAGCTATCACT 120  
Db 576 GAATTAATAACCTTTTCAGAAATATCAAGCCGCTTTTACAAGTGGCCAGCTAGCTATCACT 635  
QY 121 GAAAGACAGCAGCAATATGTTCTCGATGACACAGCAATATTTTGAGCAGATGT 180  
Db 636 GAAAGACAGCAGCAATATGTTCTCGATGACACAGCAATATTTTGAGCAGATGT 695  
QY 181 GAAGCAGCAGATGTTCCACAGAGCGCCTTACAGAAAGGATCTTCTACCGACACAGAA 240  
Db 696 GAAGCAGCAGATGTTCCACAGAGCGCCTTACAGAAAGGATCTTCTACCGACACAGAA 755  
QY 241 AAAGACACACAGCTCATCTCCAACTGTAGACTGCTGTTATGCTCGCTCGGCTGAAGATA 300  
Db 756 AAAGACACACAGCTCATCTCCAACTGTAGACTGCTGTTATGCTCGCTCGGCTGAAGATA 815  
QY 301 AGACTGACCCAGCCAGCAGCTAAGAGAAATATGCAAGTGGTCTAGCTCCACTTTA 360  
Db 816 AGACTGACCCAGCCAGCAGCTAAGAGAAATATGCAAGTGGTCTAGCTCCACTTTA 875  
QY 361 GCTTTAATAATATGTTTCATTATTTCTGCTTTTGGCTCTCTATATAAGAGCTTGT 420  
Db 876 GCTTTAATAATATGTTTCATTATTTCTGCTTTTGGCTCTCTATATAAGAGCTTGT 935  
QY 421 ATTTTCATTGGAAG 435  
Db 936 ATTTTCATTGGAAG 950  
RESULT 5  
LOCUS AX474091 651 bp DNA linear PAT 09-AUG-2002  
DEFINITION Sequence 3 from Patent WO0240691.  
ACCESSION AX474091  
VERSION AX474091.1 GI:22208240  
KEYWORDS  
SOURCE Strawberry vein banding virus (SVBV)  
ORGANISM Strawberry vein banding virus (SVBV) promoter  
Viruses; Retroid viruses; Caulimoviridae; Caulimovirus.

REFERENCE 1  
AUTHORS Wu, G. and McBride, K.  
TITLE Strawberry vein banding virus (svbv) promoter  
JOURNAL Patent: WO 0240691-A 3 23-MAY-2002;  
Maxygen, Inc. (US)  
FEATURES  
Location/Qualifiers  
1..651  
/organism="Strawberry vein banding virus"  
/mol\_type="unassigned DNA"  
/strain="E3"  
/db\_xref="taxon:47903"  
ORIGIN  
Query Match 77.5%; Score 369; DB 6; Length 651;  
Best Local Similarity 100.0%; Pred. No. 4.2e-200;  
Matches 369; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 108 GCTAGCTATCACTGAAAAAGACAGCAAGCAATGGTGTCTCGATGCAACCAACCATCT 167  
Db 278 GCTAGCTATCACTGAAAAAGACAGCAAGCAATGGTGTCTCGATGCAACCAACCATCT 337  
QY 168 TTGCACAGATGTGAAGCAGCCAGAGTGTCCACAGAGCGCACTCAGAAAAAGCATCTTC 227  
Db 338 TTGCACAGATGTGAAGCAGCCAGAGTGTCCACAGAGCGCACTCAGAAAAAGCATCTTC 397  
QY 228 TACCGACAGAAAAAGACACACAGCTCATCTCCAACTGTAGACTGTCTTATGCG 287  
Db 398 TACCGACAGAAAAAGACACACAGCTCATCTCCAACTGTAGACTGTCTTATGCG 457  
QY 288 TCGGCTGAAGATAAGACTGACCCGCGCAGCAGCACTAAGAGAAATATGCAAGTGGTCC 347  
Db 458 TCGGCTGAAGATAAGACTGACCCGCGCAGCAGCACTAAGAGAAATATGCAAGTGGTCC 517  
QY 348 TAGCTCCACTTTAGCTTTAATTAATTTATTTATTTATTTATTTCTGCTTTTCTCTAT 407  
Db 518 TAGCTCCACTTTAGCTTTAATTAATTTATTTATTTATTTCTGCTTTTCTCTAT 577  
QY 408 ATAAAGAGCTGTATTTTCAATTTGAAGCAGAGGGGAAACACACACAGAACCTCCCTGC 467  
Db 578 ATAAAGAGCTGTATTTTCAATTTGAAGCAGAGGGGAAACACACACAGAACCTCCCTGC 637  
QY 468 TTACAAACC 476  
Db 638 TTACAAACC 646  
RESULT 6  
LOCUS AX474092 701 bp DNA linear PAT 09-AUG-2002  
DEFINITION Sequence 4 from Patent WO0240691.  
ACCESSION AX474092  
VERSION AX474092.1 GI:22208241  
KEYWORDS  
SOURCE Strawberry vein banding virus (SVBV)  
ORGANISM Strawberry vein banding virus (SVBV) promoter  
Viruses; Retroid viruses; Caulimoviridae; Caulimovirus.  
REFERENCE 1  
AUTHORS Wu, G. and McBride, K.  
TITLE Strawberry vein banding virus (svbv) promoter  
JOURNAL Patent: WO 0240691-A 4 23-MAY-2002;  
Maxygen, Inc. (US)  
FEATURES  
Location/Qualifiers  
1..701  
/organism="Strawberry vein banding virus"  
/mol\_type="unassigned DNA"  
/strain="E3"  
/db\_xref="taxon:47903"  
ORIGIN  
Query Match 77.5%; Score 369; DB 6; Length 701;  
Best Local Similarity 100.0%; Pred. No. 4.2e-200;  
Matches 369; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

108 GCTAGTATCACTGAAAGACAGCAGCAATGGTGTCTCGATGACACAGAACCATCT 167  
109 |||||  
110 |||||  
111 |||||  
112 |||||  
113 |||||  
114 |||||  
115 |||||  
116 |||||  
117 |||||  
118 |||||  
119 |||||  
120 |||||  
121 |||||  
122 |||||  
123 |||||  
124 |||||  
125 |||||  
126 |||||  
127 |||||  
128 |||||  
129 |||||  
130 |||||  
131 |||||  
132 |||||  
133 |||||  
134 |||||  
135 |||||  
136 |||||  
137 |||||  
138 |||||  
139 |||||  
140 |||||  
141 |||||  
142 |||||  
143 |||||  
144 |||||  
145 |||||  
146 |||||  
147 |||||  
148 |||||  
149 |||||  
150 |||||  
151 |||||  
152 |||||  
153 |||||  
154 |||||  
155 |||||  
156 |||||  
157 |||||  
158 |||||  
159 |||||  
160 |||||  
161 |||||  
162 |||||  
163 |||||  
164 |||||  
165 |||||  
166 |||||  
167 |||||  
168 |||||  
169 |||||  
170 |||||  
171 |||||  
172 |||||  
173 |||||  
174 |||||  
175 |||||  
176 |||||  
177 |||||  
178 |||||  
179 |||||  
180 |||||  
181 |||||  
182 |||||  
183 |||||  
184 |||||  
185 |||||  
186 |||||  
187 |||||  
188 |||||  
189 |||||  
190 |||||  
191 |||||  
192 |||||  
193 |||||  
194 |||||  
195 |||||  
196 |||||  
197 |||||  
198 |||||  
199 |||||  
200 |||||  
201 |||||  
202 |||||  
203 |||||  
204 |||||  
205 |||||  
206 |||||  
207 |||||  
208 |||||  
209 |||||  
210 |||||  
211 |||||  
212 |||||  
213 |||||  
214 |||||  
215 |||||  
216 |||||  
217 |||||  
218 |||||  
219 |||||  
220 |||||  
221 |||||  
222 |||||  
223 |||||  
224 |||||  
225 |||||  
226 |||||  
227 |||||  
228 |||||  
229 |||||  
230 |||||  
231 |||||  
232 |||||  
233 |||||  
234 |||||  
235 |||||  
236 |||||  
237 |||||  
238 |||||  
239 |||||  
240 |||||  
241 |||||  
242 |||||  
243 |||||  
244 |||||  
245 |||||  
246 |||||  
247 |||||  
248 |||||  
249 |||||  
250 |||||  
251 |||||  
252 |||||  
253 |||||  
254 |||||  
255 |||||  
256 |||||  
257 |||||  
258 |||||  
259 |||||  
260 |||||  
261 |||||  
262 |||||  
263 |||||  
264 |||||  
265 |||||  
266 |||||  
267 |||||  
268 |||||  
269 |||||  
270 |||||  
271 |||||  
272 |||||  
273 |||||  
274 |||||  
275 |||||  
276 |||||  
277 |||||  
278 |||||  
279 |||||  
280 |||||  
281 |||||  
282 |||||  
283 |||||  
284 |||||  
285 |||||  
286 |||||  
287 |||||  
288 |||||  
289 |||||  
290 |||||  
291 |||||  
292 |||||  
293 |||||  
294 |||||  
295 |||||  
296 |||||  
297 |||||  
298 |||||  
299 |||||  
300 |||||  
301 |||||  
302 |||||  
303 |||||  
304 |||||  
305 |||||  
306 |||||  
307 |||||  
308 |||||  
309 |||||  
310 |||||  
311 |||||  
312 |||||  
313 |||||  
314 |||||  
315 |||||  
316 |||||  
317 |||||  
318 |||||  
319 |||||  
320 |||||  
321 |||||  
322 |||||  
323 |||||  
324 |||||  
325 |||||  
326 |||||  
327 |||||  
328 |||||  
329 |||||  
330 |||||  
331 |||||  
332 |||||  
333 |||||  
334 |||||  
335 |||||  
336 |||||  
337 |||||  
338 |||||  
339 |||||  
340 |||||  
341 |||||  
342 |||||  
343 |||||  
344 |||||  
345 |||||  
346 |||||  
347 |||||  
348 |||||  
349 |||||  
350 |||||  
351 |||||  
352 |||||  
353 |||||  
354 |||||  
355 |||||  
356 |||||  
357 |||||  
358 |||||  
359 |||||  
360 |||||  
361 |||||  
362 |||||  
363 |||||  
364 |||||  
365 |||||  
366 |||||  
367 |||||  
368 |||||  
369 |||||  
370 |||||  
371 |||||  
372 |||||  
373 |||||  
374 |||||  
375 |||||  
376 |||||  
377 |||||  
378 |||||  
379 |||||  
380 |||||  
381 |||||  
382 |||||  
383 |||||  
384 |||||  
385 |||||  
386 |||||  
387 |||||  
388 |||||  
389 |||||  
390 |||||  
391 |||||  
392 |||||  
393 |||||  
394 |||||  
395 |||||  
396 |||||  
397 |||||  
398 |||||  
399 |||||  
400 |||||  
401 |||||  
402 |||||  
403 |||||  
404 |||||  
405 |||||  
406 |||||  
407 |||||  
408 |||||  
409 |||||  
410 |||||  
411 |||||  
412 |||||  
413 |||||  
414 |||||  
415 |||||  
416 |||||  
417 |||||  
418 |||||  
419 |||||  
420 |||||  
421 |||||  
422 |||||  
423 |||||  
424 |||||  
425 |||||  
426 |||||  
427 |||||  
428 |||||  
429 |||||  
430 |||||  
431 |||||  
432 |||||  
433 |||||  
434 |||||  
435 |||||  
436 |||||  
437 |||||  
438 |||||  
439 |||||  
440 |||||  
441 |||||  
442 |||||  
443 |||||  
444 |||||  
445 |||||  
446 |||||  
447 |||||  
448 |||||  
449 |||||  
450 |||||  
451 |||||  
452 |||||  
453 |||||  
454 |||||  
455 |||||  
456 |||||  
457 |||||  
458 |||||  
459 |||||  
460 |||||  
461 |||||  
462 |||||  
463 |||||  
464 |||||  
465 |||||  
466 |||||  
467 |||||  
468 |||||  
469 |||||  
470 |||||  
471 |||||  
472 |||||  
473 |||||  
474 |||||  
475 |||||  
476 |||||  
477 |||||  
478 |||||  
479 |||||  
480 |||||  
481 |||||  
482 |||||  
483 |||||  
484 |||||  
485 |||||  
486 |||||  
487 |||||  
488 |||||  
489 |||||  
490 |||||  
491 |||||  
492 |||||  
493 |||||  
494 |||||  
495 |||||  
496 |||||  
497 |||||  
498 |||||  
499 |||||  
500 |||||  
501 |||||  
502 |||||  
503 |||||  
504 |||||  
505 |||||  
506 |||||  
507 |||||  
508 |||||  
509 |||||  
510 |||||  
511 |||||  
512 |||||  
513 |||||  
514 |||||  
515 |||||  
516 |||||  
517 |||||  
518 |||||  
519 |||||  
520 |||||  
521 |||||  
522 |||||  
523 |||||  
524 |||||  
525 |||||  
526 |||||  
527 |||||  
528 |||||  
529 |||||  
530 |||||  
531 |||||  
532 |||||  
533 |||||  
534 |||||  
535 |||||  
536 |||||  
537 |||||  
538 |||||  
539 |||||  
540 |||||  
541 |||||  
542 |||||  
543 |||||  
544 |||||  
545 |||||  
546 |||||  
547 |||||  
548 |||||  
549 |||||  
550 |||||  
551 |||||  
552 |||||  
553 |||||  
554 |||||  
555 |||||  
556 |||||  
557 |||||  
558 |||||  
559 |||||  
560 |||||  
561 |||||  
562 |||||  
563 |||||  
564 |||||  
565 |||||  
566 |||||  
567 |||||  
568 |||||  
569 |||||  
570 |||||  
571 |||||  
572 |||||  
573 |||||  
574 |||||  
575 |||||  
576 |||||  
577 |||||  
578 |||||  
579 |||||  
580 |||||  
581 |||||  
582 |||||  
583 |||||  
584 |||||  
585 |||||  
586 |||||  
587 |||||  
588 |||||  
589 |||||  
590 |||||  
591 |||||  
592 |||||  
593 |||||  
594 |||||  
595 |||||  
596 |||||  
597 |||||  
598 |||||  
599 |||||  
600 |||||  
601 |||||  
602 |||||  
603 |||||  
604 |||||  
605 |||||  
606 |||||  
607 |||||  
608 |||||  
609 |||||  
610 |||||  
611 |||||  
612 |||||  
613 |||||  
614 |||||  
615 |||||  
616 |||||  
617 |||||  
618 |||||  
619 |||||  
620 |||||  
621 |||||  
622 |||||  
623 |||||  
624 |||||  
625 |||||  
626 |||||  
627 |||||  
628 |||||  
629 |||||  
630 |||||  
631 |||||  
632 |||||  
633 |||||  
634 |||||  
635 |||||  
636 |||||  
637 |||||  
638 |||||  
639 |||||  
640 |||||  
641 |||||  
642 |||||  
643 |||||  
644 |||||  
645 |||||  
646 |||||  
647 |||||  
648 |||||  
649 |||||  
650 |||||  
651 |||||  
652 |||||  
653 |||||  
654 |||||  
655 |||||  
656 |||||  
657 |||||  
658 |||||  
659 |||||  
660 |||||  
661 |||||  
662 |||||  
663 |||||  
664 |||||  
665 |||||  
666 |||||  
667 |||||  
668 |||||  
669 |||||  
670 |||||  
671 |||||  
672 |||||  
673 |||||  
674 |||||  
675 |||||  
676 |||||  
677 |||||  
678 |||||  
679 |||||  
680 |||||  
681 |||||  
682 |||||  
683 |||||  
684 |||||  
685 |||||  
686 |||||  
687 |||||  
688 |||||  
689 |||||  
690 |||||  
691 |||||  
692 |||||  
693 |||||  
694 |||||  
695 |||||  
696 |||||  
697 |||||  
698 |||||  
699 |||||  
700 |||||  
701 |||||  
702 |||||  
703 |||||  
704 |||||  
705 |||||  
706 |||||  
707 |||||  
708 |||||  
709 |||||  
710 |||||  
711 |||||  
712 |||||  
713 |||||  
714 |||||  
715 |||||  
716 |||||  
717 |||||  
718 |||||  
719 |||||  
720 |||||  
721 |||||  
722 |||||  
723 |||||  
724 |||||  
725 |||||  
726 |||||  
727 |||||  
728 |||||  
729 |||||  
730 |||||  
731 |||||  
732 |||||  
733 |||||  
734 |||||  
735 |||||  
736 |||||  
737 |||||  
738 |||||  
739 |||||  
740 |||||  
741 |||||  
742 |||||  
743 |||||  
744 |||||  
745 |||||  
746 |||||  
747 |||||  
748 |||||  
749 |||||  
750 |||||  
751 |||||  
752 |||||  
753 |||||  
754 |||||  
755 |||||  
756 |||||  
757 |||||  
758 |||||  
759 |||||  
760 |||||  
761 |||||  
762 |||||  
763 |||||  
764 |||||  
765 |||||  
766 |||||  
767 |||||  
768 |||||  
769 |||||  
770 |||||  
771 |||||  
772 |||||  
773 |||||  
774 |||||  
775 |||||  
776 |||||  
777 |||||  
778 |||||  
779 |||||  
780 |||||  
781 |||||  
782 |||||  
783 |||||  
784 |||||  
785 |||||  
786 |||||  
787 |||||  
788 |||||  
789 |||||  
790 |||||  
791 |||||  
792 |||||  
793 |||||  
794 |||||  
795 |||||  
796 |||||  
797 |||||  
798 |||||  
799 |||||  
800 |||||  
801 |||||  
802 |||||  
803 |||||  
804 |||||  
805 |||||  
806 |||||  
807 |||||  
808 |||||  
809 |||||  
810 |||||  
811 |||||  
812 |||||  
813 |||||  
814 |||||  
815 |||||  
816 |||||  
817 |||||  
818 |||||  
819 |||||  
820 |||||  
821 |||||  
822 |||||  
823 |||||  
824 |||||  
825 |||||  
826 |||||  
827 |||||  
828 |||||  
829 |||||  
830 |||||  
831 |||||  
832 |||||  
833 |||||  
834 |||||  
835 |||||  
836 |||||  
837 |||||  
838 |||||  
839 |||||  
840 |||||  
841 |||||  
842 |||||  
843 |||||  
844 |||||  
845 |||||  
846 |||||  
847 |||||  
848 |||||  
849 |||||  
850 |||||  
851 |||||  
852 |||||  
853 |||||  
854 |||||  
855 |||||  
856 |||||  
857 |||||  
858 |||||  
859 |||||  
860 |||||  
861 |||||  
862 |||||  
863 |||||  
864 |||||  
865 |||||  
866 |||||  
867 |||||  
868 |||||  
869 |||||  
870 |||||  
871 |||||  
872 |||||  
873 |||||  
874 |||||  
875 |||||  
876 |||||  
877 |||||  
878 |||||  
879 |||||  
880 |||||  
881 |||||  
882 |||||  
883 |||||  
884 |||||  
885 |||||  
886 |||||  
887 |||||  
888 |||||  
889 |||||  
890 |||||  
891 |||||  
892 |||||  
893 |||||  
894 |||||  
895 |||||  
896 |||||  
897 |||||  
898 |||||  
899 |||||  
900 |||||  
901 |||||  
902 |||||  
903 |||||  
904 |||||  
905 |||||  
906 |||||  
907 |||||  
908 |||||  
909 |||||  
910 |||||  
911 |||||  
912 |||||  
913 |||||  
914 |||||  
915 |||||  
916 |||||  
917 |||||  
918 |||||  
919 |||||  
920 |||||  
921 |||||  
922 |||||  
923 |||||  
924 |||||  
925 |||||  
926 |||||  
927 |||||  
928 |||||  
929 |||||  
930 |||||  
931 |||||  
932 |||||  
933 |||||  
934 |||||  
935 |||||  
936 |||||  
937 |||||  
938 |||||  
939 |||||  
940 |||||  
941 |||||  
942 |||||  
943 |||||  
944 |||||  
945 |||||  
946 |||||  
947 |||||  
948 |||||  
949 |||||  
950 |||||  
951 |||||  
952 |||||  
953 |||||  
954 |||||  
955 |||||  
956 |||||  
957 |||||  
958 |||||  
959 |||||  
960 |||||  
961 |||||  
962 |||||  
963 |||||  
964 |||||  
965 |||||  
966 |||||  
967 |||||  
968 |||||  
969 |||||  
970 |||||  
971 |||||  
972 |||||  
973 |||||  
974 |||||  
975 |||||  
976 |||||  
977 |||||  
978 |||||  
979 |||||  
980 |||||  
981 |||||  
982 |||||  
983 |||||  
984 |||||  
985 |||||  
986 |||||  
987 |||||  
988 |||||  
989 |||||  
990 |||||  
991 |||||  
992 |||||  
993 |||||  
994 |||||  
995 |||||  
996 |||||  
997 |||||  
998 |||||  
999 |||||  
1000 |||||

RESULT 7  
AC005490/c  
LOCUS Homo sapiens UWGC:gl397a051 from 7p14-15, complete sequence.  
DEFINITION AC005490  
ACCESSION AC005490  
VERSION AC005490.1 GI:3420065  
KEYWORDS HTG.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE 1 (bases 1 to 41339)  
AUTHORS Iadonato,S.P., Yu,J., Wong,G.K.-S., Magness,C.L., Green,E.D., Green,P. and Olson,M.V.  
TITLE Large-scale MCD Mapping and Sequencing of Human Chromosome 7  
JOURNAL Unpublished (1998)  
REFERENCE 2 (bases 1 to 41339)  
AUTHORS Iadonato,S.P., Yu,J., Wong,G.K.-S., Magness,C.L., Green,E.D., Green,P. and Olson,M.V.  
TITLE Direct Submission  
JOURNAL Submitted (15-AUG-1998) Human Genome Center, University of Washington, Box 352145, Seattle, WA 98195, USA  
REMARK University of Washington Human Genome Center  
Box 352145 Seattle, WA 98195  
Contact: Shawn Iadonato (iadonato@u.washington.edu)  
Overlapping Sequences:  
5': UWGC:gl211a139 (Genbank Accession: AC004397)  
3': UWGC:gl397a211  
-----  
Sequence Quality Assessment:  
This entry has been annotated with sequence quality estimates computed by the Phrap assembly program.  
All manually edited bases have been reduced to quality zero.  
Quality levels above 40 are expected to have less than 1 error in 10,000 bp.  
Base-by-base quality values are not generally visible from the GenBank flat file format but are available as part of this entry's ASN.1 file.  
-----  
Double stranded (DS) coverage: 84.1%  
DS or two chemistry coverage: 99.3%  
Single stranded regions: 1  
-----  
Sequence Validation:  
-----

This sequence has been validated by Multiple Complete Digest Mapping. Comparison of the experimentally derived map digest fragments with sequence-predicted fragments is given below. Small fragments below a variable cutoff (approximately 400-600bp) are not mapped and hence do not appear in the table. There are no significant remaining discrepancies between the experimental and predicted values. Uniquely ordered fragment groups are separated by dashed lines.

Map	Seq	HindIII	Map	Seq
5297.28	5370.00	2182.20	2205.00	2532.34 2533.00
3108.32	3075.00	4569.50	4559.00	1478.00 1523.00
9618.65	9637.00	5348.89	5419.00	1272.64 1267.00
5888.88	5944.00	1113.12	1107.00	3007.80 3078.00
1303.00	1313.00	1830.87	1867.00	10170.52 10238.00
2399.52	2370.00	1723.29	1737.00	1163.00 1165.00
987.26	973.00	6076.09	6136.00	5966.04 5973.00
7906.31	7910.00	4897.48	4843.00	2022.33 2052.00
		1195.24	1206.00	2023.06 2044.00
		756.45	747.00	1012.24 1006.00
		6063.56	6050.00	

FEATURES  
source  
Location/Qualifiers  
1..41339  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="7"  
/map="7p14-15"  
/clone="NHGRI:WSS1397"  
/sub\_clone="UWGC:gl397a051"  
/cell\_lines="GM10791"  
/clone\_lib="E. Green Chromosome 7 YAC Resource"  
559..865  
/rpt\_family="Alu"  
repeat\_region  
1141..1402  
/rpt\_family="Alu"  
2148..2450  
/rpt\_family="Alu"  
4530..5123  
/rpt\_family="L1"  
repeat\_region  
complement(6377..7857)  
/rpt\_family="L1"  
8883..9178  
/rpt\_family="Alu"  
complement(12880..13092)  
/rpt\_family="MER7"  
14722..15022  
/rpt\_family="trigger2"  
15037..21430  
/rpt\_family="L1"  
16684..16908  
/rpt\_family="MER25"  
complement(21380..21587)  
/standard\_name="HSU10575"  
/note="Genbank Accession: gi|505638|gb|U10575"  
complement(21559..21855)  
/rpt\_family="MER39"  
21880..21942  
/rpt\_family="trigger2"  
complement(22410..23086)  
/rpt\_family="MER42"  
repeat\_region  
repeat\_region  
repeat\_region  
STS  
repeat\_region  
repeat\_region  
repeat\_region

```

repeat_region complement(24165. .27370)
/rpt_family="Li"
repeat_region complement(27610. .27898)
/rpt_family="Alu"
repeat_region complement(28243. .28578)
/rpt_family="L1"
/rpt_family="L1"
complement(28600. .28844)
/standard_name="HUMS2933"
/standard_name="HUMS2933"
complement(28765. .29013)
/rpt_family="MER25"
repeat_region complement(35934. .36158)
/rpt_family="MER7"
repeat_region complement(36824. .37087)
/rpt_family="Alu"
repeat_region complement(38824. .39085)
/rpt_family="Alu"
repeat_region complement(39216. .39345)
/rpt_family="MIR"

ORIGIN
Query Match 4.4% Score 21; DB 9; Length 41339;
Best Local Similarity 100.0%; Pred.No. 4;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 322 TAAGAAGAAATATGCAAGT 342
|||||
Db 26397 TAAGAAGAAATATGCAAGT 26377

RESULT 8
AC121285 192672 bp DNA linear HTG 10-MAR-2003
LOCUS Mus musculus clone RP23-437118, WORKING DRAFT SEQUENCE, 5 unordered
DEFINITION pieces.
AC121285 AC121285.3 GI:28894713
VERSION HTG: HTGS_PHASE1; HTGS_DRAFT.
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 192672)
Birren,B., Nusbaum,C. and Lander,E.
MUS musculus, clone RP23-437118
JOURNAL Unpublished
AUTHORS
* NOTE: This is a 'working draft' sequence. It currently
* consists of 5 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 23120: contig of 23120 bp in length
* 2 23221 23220: gap of 100 bp
* 3 23221 30332: contig of 7012 bp in length
* 4 30332 30333: gap of 100 bp
* 5 30333 53334: contig of 23002 bp in length
* 6 53334 53434: gap of 100 bp
* 7 53434 141535: contig of 88101 bp in length
* 8 141535 141635: gap of 100 bp
* 9 141635 192672: contig of 51037 bp in length.
* Location/Qualifiers
* 1. 192672
* /organism="Mus musculus"
* /mol_type="genomic DNA"
* /db_xref="taxon:10090"
* /clone="RP23-437118"
* /clone_lib="RPCI-23 Female Mouse BAC"
* 1. 23120
* /note="assembly_fragment"

TITLE
JOURNAL
REFERENCE 3 (bases 1 to 192672)
Submitted (16-MAY-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA

```

```

AUTHORS
Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
Anderson,S., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T.,
Boguslavsky,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y.,
Collamore,A., Cook,A., Cooke,P., Corum,B., DeArelano,K.,
Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S.,
Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S.,
Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B.,
Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
Lindblad-Toh,K., Liu,G., Lui,A., Mabbitt,R., MacLean,C.,
Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M.,
McDonald,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J.,
Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P.,
O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierce,N.,
Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P.,
Roman,J., Schauer,S., Schupback,R., Seaman,S., Severy,P., Smith,C.,
Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M.,
Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M.,
Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X.,
Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (10-MAR-2003) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Mar 10, 2003 this sequence version replaced GI:28269510.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submission@genome.wi.mit.edu
----- Project Information
Center project name: L22768
Center clone name: 437.1.18
----- Summary Statistics
Sequencing vector: Plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 190850 bases at least Q40
Consensus quality: 191537 bases at least Q30
Consensus quality: 191930 bases at least Q20
Insert size: 183000; agarose-efp
Insert size: 192272; sum-of-contigs
Quality coverage: 11.0 in Q20 bases; agarose-efp
Quality coverage: 10.5 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 5 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 23120: contig of 23120 bp in length
* 2 23221 23220: gap of 100 bp
* 3 23221 30332: contig of 7012 bp in length
* 4 30332 30333: gap of 100 bp
* 5 30333 53334: contig of 23002 bp in length
* 6 53334 53434: gap of 100 bp
* 7 53434 141535: contig of 88101 bp in length
* 8 141535 141635: gap of 100 bp
* 9 141635 192672: contig of 51037 bp in length.
* Location/Qualifiers
* 1. 192672
* /organism="Mus musculus"
* /mol_type="genomic DNA"
* /db_xref="taxon:10090"
* /clone="RP23-437118"
* /clone_lib="RPCI-23 Female Mouse BAC"
* 1. 23120
* /note="assembly_fragment"

FEATURES
source
misc_feature 1. 23120

```

```

clone_end:SP6
vector_side:left
23221..30232
/notes="assembly_fragment"
30333..53334
/notes="assembly_fragment"
53435..141535
/notes="assembly_fragment"
141636..192672
/notes="assembly_fragment"
clone_end:T7
vector_side:right"

ORIGIN
Query Match 4.4%; Score 21; DB 2; Length 192672;
Best Local Similarity 100.0%; Pred.No. 4.1;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 44 AATTAATCAAGGAGAGAAAT 64
|||||
Db 21707 AATTAATCAAGGAGAGAAAT 21727

RESULT 9
BX571704/c 221190 bp DNA linear HTG 24-SEP-2003
LOCUS
DEFINITION
Danio rerio clone DKEY-33M14, WORKING DRAFT SEQUENCE, 8 unordered
pieces.
ACCESSION BX571704
VERSION BX571704.3 GI:35209958
KEYWORDS HTG; HTGS PHASE1; HTGS DRAFT; HTGS_FULLTOP.
SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
1 (bases 1 to 221190)
McLaren S.

REFERENCE
AUTHORS
TITLE Direct Submission
JOURNAL Submitted (23-SEP-2003) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
zf1sh-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Sep 24, 2003 this sequence version replaced gi:32879563.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: zf1sh-help@sanger.ac.uk
----- Project Information
Center project name: zk33M14
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Chemistry: Dye-terminator; 100% of reads
Consensus quality: 218263 bases at least Q40
Consensus quality: 218835 bases at least Q30
Consensus quality: 219365 bases at least Q20
Insert size: 220490; sum-of-contigs
Insert size: 180894; 7.9% error; agarose-fp
Quality coverage: 6.34x in Q20 bases; sum-of-contigs Quality
coverage: 7.77x in Q20 bases; agarose-fp
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 8 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 4748: contig of 4748 bp in length
* 4749 4848: gap of 100 bp
* 4849 26640: contig of 21792 bp in length
* 26641 26740: gap of 100 bp

```

```

* 26741 39754: contig of 13014 bp in length
* 39755 39854: gap of 100 bp
* 39855 79214: contig of 39360 bp in length
* 79215 79314: gap of 100 bp
* 79315 193130: contig of 113816 bp in length
* 193131 193230: gap of 100 bp
* 193231 204142: contig of 10912 bp in length
* 204143 204243: gap of 100 bp
* 204243 207209: contig of 2967 bp in length
* 207210 207309: gap of 100 bp
* 207310 221190: contig of 13881 bp in length.

FEATURES
Location/Qualifiers
1..221190
/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/clone="DKEY-33M14"
/clone_lib="DanioKey"
1..4748
/notes="assembly_fragment:00692"
fragment_chain:1"
4849..26640
/notes="assembly_fragment:01375"
fragment_chain:1"
26741..39754
/notes="assembly_fragment:01249"
fragment_chain:1"
39855..79214
/notes="assembly_fragment:00365"
79315..193130
/notes="assembly_fragment:00635"
193231..204142
/notes="assembly_fragment:00668"
204243..207209
/notes="assembly_fragment:00870.0"
207310..221190
/notes="assembly_fragment:00990"

ORIGIN
Query Match 4.4%; Score 21; DB 2; Length 221190;
Best Local Similarity 100.0%; Pred.No. 4.2;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 363 TTTAATAATTATGTTTCATTA 383
|||||
Db 69946 TTTAATAATTATGTTTCATTA 69926

RESULT 10
AK061232
LOCUS
DEFINITION
Oryza sativa (japonica cultivar-group) cDNA clone:006-211-B02, full
insert sequence.
ACCESSION AK061232
VERSION AK061232.1 GI:32971250
KEYWORDS FLI_CDNA; oligo-capping.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
1

REFERENCE
AUTHORS
The Rice Full-Length cDNA Consortium, National Institute of
Agrobiological Sciences Rice Full-Length cDNA Project Team;
Kikuchi,S., Satoh,K., Nagata,T., Kawagashira,N., Doi,K.,
Kishimoto,N., Yazaki,J., Ishikawa,M., Yamada,H., Ooka,H., Hotta,I.,
Kojima,K., Namiki,T., Ohneda,E., Yahagi,W., Suzuki,K., Li,C.,
Ohtsuki,K., Shishiki,T., Foundation of Advancement of International
Science Genome Sequencing & Analysis Group; Otsomo,Y., Murakami,K.,
Iida,Y., Sugano,S., Fujimura,T., Suzuki,Y., Tsunoda,Y.,
Kurosaki,T., Kodama,T., Masuda,H., Kobayashi,M., Xie,Q., Lu,M.,
Narikawa,R., Sugiyama,A., Mizuno,K., Yokomizo,S., Nikura,J.,
Ikeda,R., Ishibiki,J., Kawamata,M., Yoshimura,A., Miura,J.,

```

Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN:  
 Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S.,  
 Hara, A., Hashidume, W., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M.,  
 Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Osato, N., Oka, Y.,  
 Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,  
 Yoshino, M., and Hayashizaki, Y.  
 Collection, mapping, and annotation of over 28,000 cDNA clones from  
 japonica rice

Science 301 (5631), 376-379 (2003)

22752273

12869764

2 (bases 1 to 1046)

ADACHI, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K.,  
 Fujiwara, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W.,  
 Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hiroaka, T.,  
 Hori, F., Hotta, I., Iida, J., Iida, Y., Ikeda, R., Imamura, K.,  
 Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I.,  
 Kanagawa, S., Katoh, H., Kawagashira, N., Kawai, J., Kawamata, M.,  
 Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M.,  
 Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,  
 Kova, S., Kurihara, C., Kurosaki, T., Kusumegi, T., Li, C., Lu, M.,  
 Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A.,  
 Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, M.,  
 Namiki, T., Narikawa, R., Niikura, J., Nishi, K., Nomura, K.,  
 Numasaki, R., Ohneda, E., Ohno, M., Ohtsuki, K., Oka, M., Ooka, H.,  
 Osato, N., Ota, Y., Otsu, Y., Ryu, R., Saitoh, H., Sakai, C., Sakai, K.,  
 Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K.,  
 Shinagawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S.,  
 Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y.,  
 Tagawa, A., Tanahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A.,  
 Toyota, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yanaghi, W.,  
 Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S., and  
 Yoshimura, A.

Direct Submission

Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of  
 Agrobiological Sciences, Department of Molecular Genetics, Head of  
 Laboratory of Gene Expression, 2-1-2 Kannondai, Tsukuba, Ibaraki  
 305-8602, Japan (E-mail: skikuchi@nias.affrc.go.jp,  
 Tel: 81-29-838-7007, Fax: 81-29-838-7007)

This clone is one of the 28K full-length cDNA clones from japonica  
 rice.

URL: <http://cdna01.dna.affrc.go.jp/cdna/>

NIAS Rice Full-Length cDNA Project Team: Kikuchi, S., Satoh, K.,  
 Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J.,  
 Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T.,  
 Ohneda, E., Yanaghi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T., and  
 Yamamoto, M.

FAIS Genome Sequencing & Analysis Group: Otsu, Y., Iida, Y.,  
 Fujiwara, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M.,  
 Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J.,  
 Mizuno, K., Narikawa, R., Niikura, J., Oka, M., Sano, H., Sugano, S.,  
 Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S.,  
 Yoshimura, A., Matsubara, K., and Murakami, K.  
 Genome Exploration Research Group in Riken Genomic Sciences Center  
 and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K.,  
 Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T.,  
 Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K.,  
 Hiraoka, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y.,  
 Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J.,  
 Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,  
 Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,  
 Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Osato, N.,  
 Ota, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H.,  
 Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,  
 Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Tanahashi, F.,  
 Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toyota, T., Waki, K.,  
 Yasunishi, A., and Hayashizaki, Y.

Location/Qualifiers

1..1046  
 /organism="Oryza sativa (japonica cultivar-group)"  
 /mol\_type="mRNA"  
 /cultiVar="Nipponbare"  
 /db\_xref="taxon:39947"

/clone="006-211-B02"

ORIGIN  
 Query Match 4.2%; Score 20; DB 8; Length 1046;  
 Best Local Similarity 100.0%; Pred. No. 13;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 413 GAGCTTGATTTTCATTGCA 432

|||||

Db 465 GAGCTTGATTTTCATTGCA 484

RESULT 11

SCYKL041W

LOCUS

SCYKL041W

DEFINITION

S.cerevisiae chromosome XI reading frame ORF YKL041w.

ACCESSION

Z28041.Y13137

VERSION

Z28041.1 GI:486051

KEYWORDS

SOURCE

Saccharomyces cerevisiae (baker's yeast)

ORGANISM

Saccharomyces cerevisiae

Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

Saccharomycetales; Saccharomycetaceae; Saccharomycetes.

REFERENCE

1 (bases 1 to 1490)

AUTHORS

Furnelle, B., Skala, J., van Dyck, L., Tettelin, H. and Goffeau, A.

JOURNAL

Unpublished

REFERENCE

2 (bases 1 to 1490)

AUTHORS

MIPS

Direct Submission

Submitted (09-MAY-1994) Data collected by MIPS on behalf of the

European yeast chromosome XI sequencing project. MIPS at the

Max-Planck-Institut fuer Biochemie, Am Klopferspitz 18a D-82152

Martinsried, FRG; E-mail: Mewes@hpmpic.mips.biochem.mpg.de

FEATURES

Location/Qualifiers

1..1490

/organism="Saccharomyces cerevisiae"

/mol\_type="genomic DNA"

/db\_xref="taxon:4932"

/chromosome="XI"

576..1250

/gene="VPS24"

576..1250

/gene="VPS24"

/note="ORF YKL041w"

/codon\_start=1

/protein\_id="CAA81876.1"

/db\_xref="GI:486052"

/db\_xref="GOA:P36095"

/db\_xref="SGD:S0001524"

/db\_xref="SWISS-PROT:P36095"

/translation="MDYIKKAIWGDDPKQQRRTSVLRKNGRIEKLRLTLVQLNK

TQQLIKKAKNDVRLIYAKELQINKQDMYTSRAQLDSVRMKIDAIRNLTLS

NQMDSAGLMREVNLRPLQRLNTMIELEKELKSGIISENVDDTMSVGDVGDEMD

EAVDDEVKNIVEQYTNKRFKNVDQVPTVELAANEEEOIPDEKVDDEADRMVNERER

LRALQN"

ORIGIN

Query Match 4.2%; Score 20; DB 8; Length 1490;

Best Local Similarity 100.0%; Pred. No. 13;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 416 CTGTATTTTCATTGCAAGG 435

|||||

Db 1462 CTGTATTTTCATTGCAAGG 1481

RESULT 12

AX008583

LOCUS

AX008583

DEFINITION

Sequence 1 from Patent WO9966056.

ACCESSION

AX008583

VERSION

AX008583.1 GI:9996133

KEYWORDS

FEATURES

source

AX008583

DEFINITION

Sequence 1 from Patent WO9966056.

ACCESSION

AX008583

VERSION

AX008583.1 GI:9996133

KEYWORDS

AX008583

LOCUS

AX008583

DEFINITION

Sequence 1 from Patent WO9966056.

ACCESSION

AX008583

VERSION

AX008583.1 GI:9996133

KEYWORDS

AX008583

LOCUS

AX008583

DEFINITION

Sequence 1 from Patent WO9966056.

ACCESSION

AX008583

VERSION

AX008583.1 GI:9996133

KEYWORDS



```

SOURCE      Chlamydomonas reinhardtii
ORGANISM    Chlamydomonas reinhardtii
            Chlamydomonadaceae; Chlorophyta; Chlorophyceae; Volvocales;
REFERENCE 1
AUTHORS    Ball,S.
TITLE      Method for obtaining modified polysaccharides
JOURNAL    Patent: WO 996056-A 1 23-DEC-1999;
            BALL STEVEN (FR); BIOENMA (FR)
FEATURES   Location/Qualifiers
            source
              1..1557
                /organism="Chlamydomonas reinhardtii"
                /mol_type="unassigned DNA"
                /db_xref="taxon:3055"
ORIGIN
Query Match      4.2%; Score 20; DB 6; Length 1557;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 87 AGCCCGCTTTACAAGTGCC 106
      |||||||||||||||||||
DB 1441 AGCCCGCTTTACAAGTGCC 1460

RESULT 13
LOCUS      SCYKL040C
DEFINITION S.cerevisiae chromosome XI reading frame ORF YKL040c.
ACCESSION  Z28040 Y13137
VERSION     Z28040.1 GI:486049
KEYWORDS
SOURCE      Saccharomyces cerevisiae (baker's yeast)
ORGANISM    Saccharomyces cerevisiae
            Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
            Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
REFERENCE 1 (bases 1 to 1804)
AUTHORS    Purnelle,B., Skala,J., van Dyck,L., Tettelin,H. and Goffeau,A.
JOURNAL    Unpublished
REFERENCE 2 (bases 1 to 1804)
AUTHORS    MIPS.
TITLE      Direct Submission
JOURNAL    Submitted (09-MAY-1994) Data collected by MIPS on behalf of the
            European yeast chromosome XI sequencing project. MIPS at the
            Max-Planck-Institut fuer Biochemie, Am Klopferspitz 18a D-82152
            Martinsried, FRG; E-mail: Mewese@hgm.mc.mpg.de
FEATURES   Location/Qualifiers
            source
              1..1804
                /organism="Saccharomyces cerevisiae"
                /mol_type="genomic DNA"
                /db_xref="taxon:4932"
                /chromosome="XI"
                /complement(241..1011)
                /note="unnamed protein product; ORF YKL040c"
                /codon_start=1
                /protein_id="CAA81875.1"
                /db_xref="GI:486050"
                /db_xref="GOA:P32860"
                /db_xref="SGD:S0001523"
                /db_xref="SWISS-PROT:P32860"
                /translation="MFKSVAKLGKSPFYLNQSRILIKTLTPNENALKFLSTDGEM
            LQTRGSKIIVKNTDENLINSKLAQOIPLQCPVESLMIGDDFLINKDRMVHNSI
            KPFIIDLTLKANGEDVVISKEPHAVQEEGGYKINPKFELTBEDEVSELIJEL
            IDIRPAILEDGGDIDYRWDPKTVYLRLOGACTSCSSSEVTIKYGIESMLKHVY
            DEKVEIQIMDPQEIALKFEDKLEKLESSKNTSHEK"
CDS
ORIGIN
Query Match      4.2%; Score 20; DB 8; Length 1804;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 416 CTTGATTTTCATTGAAGG 435
      |||||||||||||||||||

```

```

DB 212 CTTGATTTTCATTGAAGG 231

RESULT 14
LOCUS      AF307842
DEFINITION Chlamydomonas reinhardtii 4-alpha-glucanotransferase (STAl1) mRNA,
            complete cds.
ACCESSION  AF307842
VERSION     AF307842.1 GI:11095334
KEYWORDS
SOURCE      Chlamydomonas reinhardtii
ORGANISM    Chlamydomonas reinhardtii
            Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
            Chlamydomonadaceae; Chlamydomonas.
REFERENCE 1 (bases 1 to 2204)
AUTHORS    Watebled,F., Rai,J.P., Dauvillee,D., Myers,A.M., James,M.G.,
            Schlichting,R., Giersch,C., Ball,S.G. and D'Hulst,C.
TITLE      STAl1, a Chlamydomonas reinhardtii locus required for Normal Starch
            Granule Biogenesis, Encodes Disproportionating Enzyme. Further
            Evidence for a Function of alpha-1,4 Glucanotransferases during
            Starch Granule Biosynthesis in Green Algae
JOURNAL    Plant Physiol. 132 (1), 137-145 (2003)
PUBMED     12746519
REFERENCE 2 (bases 1 to 2204)
AUTHORS    Watebled,F., D'Hulst,C., Myers,A.M. and Ball,S.G.
TITLE      Direct Submission
JOURNAL    Submitted (22-SEP-2000) UMR8576 of CNRS, Laboratory of Biological
            Chemistry, University of Sciences and Technologies of Lille,
            Villeneuve d'Ascq Cedex 59655, France
FEATURES   Location/Qualifiers
            source
              1..2204
                /organism="Chlamydomonas reinhardtii"
                /mol_type="mRNA"
                /db_xref="taxon:3055"
              1..2204
                /gene="STAl1"
                /gene="STAl1"
                /note="disproportionating enzyme; D-enzyme"
                /citation=[1]
                /codon_start=1
                /product="4-alpha-glucanotransferase"
                /protein_id="AAG29839.1"
                /db_xref="GI:11095335"
                /translation="WIPLRAPQGVRYTQIDRSQVLPQLRSRPAQSVSHVAAAGSRRRA
            VSRSVRAFAWGTPTKTPGAPGVLDKDYDQKIPSYNKRERAGILLHPTSLPGPYGIG
            DIGDEAKRFVDMLADHGMQWQLPLVPDPMTYSPSYGTDCANGCNPLVSIIELIKD
            GLIFSETPRPVITADVDYPAVAAKLLPLKRAAQRLLKEDRFLRLEEYLYKREKHP
            WVEDSALFDVARNLPQLSOLAWMWPSPRLRQXKALKERETNKDAIDRFVVIQYFF
            EKQWKAIRSYANGKIGKILGDMPIYVGSHSADWANSHPLENEAGLPEQYSGVPPDA
            FSATGQLWGSPLYKWPAAHKSGFKWWTARMTARLTLELDECRIDHFRFAGYWSVDANE
            TTAMGNWRQFGLELFTAMKALGAVPIAEIDLGVITTDVVALREAGIAGPMVVLQF
            AWGGPGNVHLPHNHYENCYFPGTHDNETAVGWRFGSANDTKSYIKSLRTDGGDI
            AWDFTACMAAVPRTCVTQMVDVRLONTARMTPTGTAAGNWRMRMGDGNVWNSLKKE
            GEDLRKVAHDTNRLPKPKA"
ORIGIN
Query Match      4.2%; Score 20; DB 8; Length 2204;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 87 AGCCCGCTTTACAAGTGCC 106
      |||||||||||||||||||
DB 1159 AGCCCGCTTTACAAGTGCC 1178

RESULT 15
LOCUS      BC051144/c
DEFINITION Mus musculus RIKEN cDNA C330007P06 gene, mRNA (cDNA clone MGC:58463
            IMAGE:6589980), complete cds.
ACCESSION  BC051144

```

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

REMARK

COMMENT

BC051144.1

GI:30047895

MGC.

Mus musculus (house mouse)

Mus musculus

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 2709)

Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D., Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Diachenko,L., Marusina,K., Farmer,A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Schmitt,E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raha,S.S., Lequellano,N.A., Peters,G.J., Abramson,R.D., Mullaly,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Woxley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahay,J., Helton,E., Kettelman,M., Madan,A., Rodrigues,S., Sanchez,A., Whitting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E., Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

22388257

12477932

2 (bases 1 to 2709)

Strausberg,R.

Direct Submission

Submitted (11-APR-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)

Tissue Procurement: Dr. Michael Brownstein

cDNA Library Preparation: Michael Brownstein / Ted Usdin

Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Genome Sequence Centre, BC Cancer Agency, Vancouver, BC, Canada

[info@bcgsc.bc.ca](mailto:info@bcgsc.bc.ca)

Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield, Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin, Letticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline Schein, Duane Smailus, Michael Smith, Lorraine Spence, Jeff Scott, Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Marra.

FEATURES

source

1. .2709

/organism="Mus musculus"

/mol\_type="mRNA"

/db\_xref="taxon:10090"

/clone="MGC:58463 IMAGE:6588980"

/tissue\_type="Brain, mouse"

/clone\_lib="NIH\_MGC\_144"

/lab\_host="DH10B"

/note="vector: pDNR-LIB"

1. .2709

gene

ORIGIN

Query Match 4.2%; Score 20; DB 10; Length 2709;

Best Local Similarity 100.0%; Pred. No. 14;

Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy

369 AATTATGTTTCATTATTATT 388

|||||

Db

2398 AATTATGTTTCATTATTATT 2379

|||||

Search completed: July 31, 2004, 10:16:38

Job time : 2310 secs

1

28..596

/codon\_start=1

/product="C330007P06Rik protein"

/db\_xref="LocusID:77644"

/db\_xref="MGI:1924894"

/protein\_id="AAH51144.1"

/db\_xref="GI:30047896"

/db\_xref="LocusID:77644"

/translation="MPKVVSRSVCSGSDTDEEYDDGKPLHVVYCLGQMVLDCCQLEKLPWRPDRSRVIDRAAKHAHKFNTDEETTYLEREGIEROYKKCAKGLPLFYQSQPNAPVTFTVDGAVVKFGCGFKNTIYTKQBPFPKVMYTKTKDMGKPSSTVSTIDEEEBEIEAREVADSYAQNKVLEKQLERKMSKRRLLQELAELEAKKAKMKGTLLIDNQFK"

CDS

28..596

/gene="C330007P06Rik"

/db\_xref="LocusID:77644"

/db\_xref="MGI:1924894"

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 31, 2004, 06:52:19 ; Search time 307 Seconds  
(without alignments)  
6586.783 Million cell updates/sec

Title: US-10-020-540A-1

Perfect score: 476

Sequence: 1 aactatgctgacgaagat.....aacctccctgcttacaacc 476

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_29Jan04.\*

- 1: Geneseq1980s.\*
- 2: Geneseq1990s.\*
- 3: Geneseq2000s.\*
- 4: Geneseq2001as.\*
- 5: Geneseq2001bs.\*
- 6: Geneseq2002s.\*
- 7: Geneseq2003as.\*
- 8: Geneseq2003bs.\*
- 9: Geneseq2003cs.\*
- 10: Geneseq2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	476	100.0	476	6	AAD39132 Strawberry
2	476	100.0	518	6	AAD39133 Strawberry
3	476	100.0	1057	6	ABX09816 Promoter
4	476	100.0	1071	6	ABX09826 SVBV type
5	369	77.5	651	6	AAD39134 Strawberry
6	369	77.5	701	6	AAD39135 Strawberry
7	37.6	7.9	3294	4	ABL26902 Drosophila
8	37.6	7.9	78845	3	AAAG1463 Continuation (9 of
9	37.6	7.9	110000	3	AAAG1490 Continuation (4 of
10	37.6	7.9	349980	3	AAAF1608 Continuation (9 of
11	36.8	7.7	6012	4	AAS46724 Tumour su
12	36.6	7.7	17211	6	ABL32655 Human imm
13	36.4	7.6	285	7	ABZ38820 N. gonorr
14	36.2	7.6	110000	7	AAD32224_2 Continuation (3 of
15	36.2	7.6	110000	7	AAD53224_3 Continuation (4 of
16	35.8	7.5	963	3	AAA70197 Plasmodiu
17	35.2	7.4	160	5	ABV56034 Human pro
18	35.2	7.4	8992	6	ABK31421 Signal tr
19	35	7.4	458	4	AAS28807 Human imm
20	35	7.4	458	4	AAS27944 Novel cdn
21	35	7.4	458	4	ABA06606 Human cdn
22	35	7.4	458	6	ABV83943 Human pol
23	35	7.4	458	9	ADB31532 Human cdn

24	35	7.4	502	4	ABA06349 Human CDN
25	35	7.4	502	6	ABV83686 Human pol
26	35	7.4	504	4	AAS28903 Human imm
27	35	7.4	504	4	AAS28336 Genomic s
28	35	7.4	504	9	ADB31744 Human nov
29	34.8	7.3	112190	4	AAD44801 Human GPC
30	34.6	7.3	923	4	AAD05641 Human sec
31	34.6	7.3	923	4	AAD08435 Human sec
32	34.6	7.3	923	7	ADA56482 Gene enco
33	34.6	7.3	923	7	ADA56482 Human sec
34	34.6	7.3	923	7	ADA40320 Human sec
35	34.6	7.3	923	9	ADK73842 Human sec
36	34.6	7.3	6314	4	AAK90962 Human dig
37	34.6	7.3	6314	5	AAS31997 Human liv
38	34.6	7.3	6314	6	ABN90352 Human liv
39	34.6	7.3	9289	4	AAK90960 Human dig
40	34.6	7.3	9289	5	AAS31995 Human liv
41	34.6	7.3	9289	6	ABN90350 Human liv
42	34.6	7.3	17865	4	AAK90961 Human dig
43	34.6	7.3	17865	5	AAS31996 Human liv
44	34.6	7.3	49047	8	ABN90351 Human liv
45	34.6	7.3	49047	9	ADA02633 Mouse Dnt
					ADB72371 Mouse Dnt

ALIGNMENTS

RESULT 1

AAD39132	standard; DNA; 476 BP.
XX	
AC	AAD39132;
XX	
DT	04-OCT-2002 (first entry)
XX	
DE	Strawberry vein banding virus (SVBV) promoter fragment DNA.
XX	
KW	Strawberry vein banding virus promoter; transgenic plant; transgenic; ds.
XX	
OS	Strawberry vein banding virus.
XX	
PN	WO200240691-A2.
XX	
PD	23-MAY-2002.
XX	
PF	30-OCT-2001; 2001WO-US047964.
XX	
PR	01-NOV-2000; 2000US-0245354P.
XX	
PA	(MAXY-) MAXYGEN INC.
XX	
PI	Wu G, McBride K;
XX	
DR	WPI; 2002-479908/51.
XX	
PT	Novel isolated or recombinant nucleic acid comprising strawberry vein banding virus promoter operably linked to a heterologous polynucleotide, useful for expressing the heterologous polynucleotide in a plant cell.
PT	
XX	Claim 1; Page 47; 48pp; English.
PS	
XX	The invention relates to an isolated/recombinant nucleic acid comprising a strawberry vein banding virus (SVBV) promoter operably linked to a heterologous polynucleotide. The invention is useful for expressing a heterologous polynucleotide in a plant cell, by introducing the nucleic acid into a plant cell, where the plant cell is present within a plant cell. The invention is used to introduce the isolated nucleic acid into the cell. The invention is useful to drive gene expression in plant cells and transgenic plants, for transforming plant cells and producing transgenic plants, for compensating for missing or altered gene expression in a plant, for expressing an endogenous protein at higher than normal levels, for expressing a novel gene in a plant, and to suppress expression of endogenous plant genes. The present sequence is SVBV-E3 promoter fragment

CC DNA  
XX  
SQ Sequence 476 BP; 164 A; 109 C; 87 G; 116 T; 0 U; 0 Other;  
Query Match 100.0%; Score 476; DB 6; Length 476;  
Best Local Similarity 100.0%; Pred. No. 1.8e-123;  
Matches 476; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 AACTATGCTGATGACAGATTAATCTTAATAGCAATTAATTCAGAAATTAATCAAGGAGAAA 60  
Db 1 AACTATGCTGATGACAGATTAATCTTAATAGCAATTAATTCAGAAATTAATCAAGGAGAAA 60  
QY 61 GAATTAATAAATCTTTTTCAGATATGAAGCCGCTTTTCAAGTGGCCAGCTAGCTATCACT 120  
Db 61 GAATTAATAAATCTTTTTCAGATATGAAGCCGCTTTTCAAGTGGCCAGCTAGCTATCACT 120  
QY 121 GAAAGACAGCAAGCAATGCTGCTGATGACACAGAACACATCTTTGCGAGCATGT 180  
Db 121 GAAAGACAGCAAGCAATGCTGCTGATGACACAGAACACATCTTTGCGAGCATGT 180  
QY 181 GAAGCAGCAGAGTGGTCCCAAGACGCACTCAGAAAGGCATCTTTACCGACACAGAA 240  
Db 181 GAAGCAGCAGAGTGGTCCCAAGACGCACTCAGAAAGGCATCTTTACCGACACAGAA 240  
QY 241 AAGACACACACAGCTCATCTCAACATGTAGACTGCTGTTATGCGTGGCTGAAGATA 300  
Db 241 AAGACACACACAGCTCATCTCAACATGTAGACTGCTGTTATGCGTGGCTGAAGATA 300  
QY 301 AGACTGACCCCGAGGCGCAGCACTAAGAGAGAAATTAATGCAAGTGGTCTAGCTCCACTTTA 360  
Db 301 AGACTGACCCCGAGGCGCAGCACTAAGAGAGAAATTAATGCAAGTGGTCTAGCTCCACTTTA 360  
QY 361 GCTTTAATAATATGTTTTCATATATATCTCTGCTTTTGTCTCTCTATATAAGAGCTTGT 420  
Db 361 GCTTTAATAATATGTTTTCATATATATCTCTGCTTTTGTCTCTCTATATAAGAGCTTGT 420  
QY 421 ATTTTCATTTGAAGCGAGCGGCAACACACACAGAACCTCCCTGCTTACAAACC 476  
Db 421 ATTTTCATTTGAAGCGAGCGGCAACACACACAGAACCTCCCTGCTTACAAACC 476

RESULT 2  
AAD39133  
ID AAD39133 standard; DNA; 518 BP.  
XX  
AC AAD39133;  
XX  
DT 04-OCT-2002 (first entry)  
XX  
DE Strawberry vein banding virus (SVBV) promoter DNA.  
XX  
KW Strawberry vein banding virus promoter; transgenic plant; transgenic; ds.  
XX  
OS Strawberry vein banding virus.  
XX  
FN WO200240691-A2.  
XX  
PD 23-MAY-2002.  
XX  
PF 30-OCT-2001; 2001WO-US047964.  
XX  
PR 01-NOV-2000; 2000US-0245354P.  
XX  
PA (MAXY-) MAXYGEN INC.  
XX  
PI Wu G, McBride K;  
XX  
DR WPI; 2002-479908/51.  
XX  
PT Novel isolated or recombinant nucleic acid comprising strawberry vein  
banding virus promoter operably linked to a heterologous polynucleotide,  
PT useful for expressing the heterologous polynucleotide in a plant cell.

PS Claim 1; Page 47; 48pp; English.  
XX  
CC The invention relates to an isolated/recombinant nucleic acid comprising  
a strawberry vein banding virus (SVBV) promoter operably linked to a  
heterologous polynucleotide. The invention is useful for expressing a  
heterologous polynucleotide in a plant cell, by introducing the nucleic  
acid into a plant cell, where the plant cell is present within a plant  
cell. The invention is used to introduce the isolated nucleic acid into the  
cell. Agrobacterium is used to introduce the isolated nucleic acid into the  
cell. The invention is useful to drive gene expression in plant cells and  
transgenic plants, for transforming plant cells and producing transgenic  
plants, for compensating for missing or altered gene expression in a  
plant, for expressing an endogenous protein at higher than normal levels,  
for expressing a novel gene in a plant, and to suppress expression of  
endogenous plant genes. The present sequence is SVBV promoter DNA  
XX  
SQ Sequence 518 BP; 171 A; 122 C; 100 G; 125 T; 0 U; 0 Other;  
Query Match 100.0%; Score 476; DB 6; Length 518;  
Best Local Similarity 100.0%; Pred. No. 1.8e-123;  
Matches 476; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 AACTATGCTGATGACAGATTAATCTTAATAGCAATTAATTCAGAAATTAATCAAGGAGAAA 60  
Db 15 AACTATGCTGATGACAGATTAATCTTAATAGCAATTAATTCAGAAATTAATCAAGGAGAAA 74  
QY 61 GAATTAATAAATCTTTTTCAGATATGAAGCCGCTTTTCAAGTGGCCAGCTAGCTATCACT 120  
Db 75 GAATTAATAAATCTTTTTCAGATATGAAGCCGCTTTTCAAGTGGCCAGCTAGCTATCACT 134  
QY 121 GAAAGACAGCAAGCAATGCTGCTGATGACACAGAACACATCTTTGCGAGCATGT 180  
Db 135 GAAAGACAGCAAGCAATGCTGCTGATGACACAGAACACATCTTTGCGAGCATGT 194  
QY 181 GAAGCAGCAGAGTGGTCCCAAGACGCACTCAGAAAGGCATCTTTACCGACACAGAA 240  
Db 195 GAAGCAGCAGAGTGGTCCCAAGACGCACTCAGAAAGGCATCTTTACCGACACAGAA 254  
QY 241 AAGACACACACAGCTCATCTCAACATGTAGACTGCTGTTATGCGTGGCTGAAGATA 300  
Db 255 AAGACACACACAGCTCATCTCAACATGTAGACTGCTGTTATGCGTGGCTGAAGATA 314  
QY 301 AGACTGACCCCGAGGCGCAGCACTAAGAGAGAAATTAATGCAAGTGGTCTAGCTCCACTTTA 360  
Db 315 AGACTGACCCCGAGGCGCAGCACTAAGAGAGAAATTAATGCAAGTGGTCTAGCTCCACTTTA 374  
QY 361 GCTTTAATAATATGTTTTCATATATATCTCTGCTTTTGTCTCTCTATATAAGAGCTTGT 420  
Db 375 GCTTTAATAATATGTTTTCATATATATCTCTGCTTTTGTCTCTCTATATAAGAGCTTGT 434  
QY 421 ATTTTCATTTGAAGCGAGCGGCAACACACACAGAACCTCCCTGCTTACAAACC 476  
Db 435 ATTTTCATTTGAAGCGAGCGGCAACACACACAGAACCTCCCTGCTTACAAACC 490

RESULT 3  
ABX09816  
ID ABX09816 standard; DNA; 1057 BP.  
XX  
AC ABX09816;  
XX  
DT 22-JAN-2003 (first entry)  
XX  
DE Promoter sequence #1 useful in method for nematode control.  
XX  
KW RNAi molecule; double-stranded interfering RNA; nematode control;  
KW RNA mediated interference; mRNA transcript; nematode gene; growth;  
KW development; parasitism; reproduction; RNAi vector; mRNA translation;  
KW nematode inhibitor; agricultural industry; anti-nematode; promoter; ds.  
XX  
OS Unidentified.  
XX  
PN WO200196584-A2.  
XX

PD 20-DEC-2001.  
XX  
PF 12-JUN-2001; 2001WO-US018911.  
XX  
PR 12-JUN-2000; 2000US-0210917P.  
XX  
PA (AKKA-) AKKADIX CORP.  
XX  
PI Mushegian AR, Taylor CG, Feitelson JS, Broshkin AM;  
XX WPI; 2002-139714/18.  
DR  
XX  
XX RNA mediated interference molecule useful for disrupting cellular process  
PT in a nematode, for controlling nematodes comprises genetic regulatory  
PT sequences.  
XX  
PS Claim 143; Page 44; 103pp; English.  
XX  
XX The present invention relates to RNAi (double-stranded interfering RNA or  
CC RNA mediated interference) molecules (nematode genes), and methods of  
CC using these sequences in nematode control. RNAi molecules selectively  
CC target mRNA transcripts of essential nematode genes. The RNAi molecules  
CC of the invention are useful for disrupting cellular processes in a  
CC nematode by contacting the nematode with a composition comprising an RNAi  
CC molecule. The RNAi molecules are useful for killing nematodes and/or  
CC inhibiting their growth, development, parasitism or reproduction and also  
CC for the regulation of levels of specific mRNA in nematodes. Genetic  
CC regulatory sequences such as promoters, enhancers and terminators can be  
CC used in genetic constructs such as RNAi vectors which can be used for  
CC nematode control. The RNAi molecules are capable of targeting and  
CC reducing (and, in some cases, preventing) the translation of a specific  
CC gene product, and can be used to reduce or prevent mRNA translation in  
CC any tissue of the nematode because of its ability to cross tissue and  
CC cellular boundaries. The RNAi molecule can be contacted with a nematode  
CC by soaking, injection, or consumption of a food source containing an RNAi  
CC molecule. The RNAi molecules can also be used as an epigenetic factor to  
CC prevent the proliferation of subsequent generations of nematodes, to  
CC produce nematode inhibitors or RNAi in the plants, and provide new  
CC biotechnological strategies for managing nematodes under sustainable  
CC agricultural conditions. ABX09816-ABX09823 represent promoter sequences  
XX useful in genetic constructs for the control of nematodes  
XX  
SQ Sequence 1057 BP; 367 A; 230 C; 200 G; 260 T; 0 U; 0 Other;  
  
Query Match 100.0%; Score 476; DB 6; Length 1057;  
Best Local Similarity 100.0%; Pred. NO. 2.4e-123;  
Matches 476; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1 AACTATGCTGATGACAGATAATTTCTAATAGCAATTTTCAAGATTAATCAGAGGAGAA 60  
Db |||||  
555 AACTATGCTGATGACAGATAATTTCTAATAGCAATTTTCAAGATTAATCAGAGGAGAA 614  
Qy 61 GAATTAATACTCTTTTTCAGATATGAGCCCGCTTTTACAGTGGCCAGCTAGTATCACT 120  
Db |||||  
615 GAATTAATACTCTTTTTCAGATATGAGCCCGCTTTTACAGTGGCCAGCTAGTATCACT 674  
Qy 121 GAAGACAGCAGACATATGCTGCTGATGACACAGACACATCTTTTCACAGATGT 180  
Db |||||  
675 GAAGACAGCAGACATATGCTGCTGATGACACAGACACATCTTTTCACAGATGT 734  
Qy 181 GAAGACAGCAGATGGTCCACAGACACACCTCTTACAGAAAGCATCTTTACCCACAGAA 240  
Db |||||  
735 GAAGACAGCAGATGGTCCACAGACACACCTCTTACAGAAAGCATCTTTACCCACAGAA 794  
Qy 241 AAAGACACACAGCTCATCATCAATCAATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300  
Db |||||  
795 AAAGACACACAGCTCATCATCAATCAATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 854  
Qy 301 AGACTGACCCAGCCAGCAGCTAAAGAGAAATAATGCAAGTGGTCTTACTCCACTTTA 360  
Db |||||  
855 AGACTGACCCAGCCAGCAGCTAAAGAGAAATAATGCAAGTGGTCTTACTCCACTTTA 914  
Qy 361 GCTTTAATAATATGTTTTCATTAATTTCTCTGCTTTTCTCTCTATATAAGAGCTTGT 420

Db 915 GCTTTAATAATATGTTTTCATTAATTTCTCTGCTTTTCTCTATATAAGAGCTTGT 974  
Qy 421 ATTTTCATTTGAAGCGAGGCGGAACACACACAGAACCTCTCTGTTTACAAACC 476  
Db 975 ATTTTCATTTGAAGCGAGGCGGAACACACACAGAACCTCTCTGTTTACAAACC 1030  
  
RESULT 4  
ABX09826  
ID ABX09826 standard; DNA; 1071 BP.  
XX  
XX AC ABX09826;  
XX  
DT 29-AUG-2003 (revised)  
DT 22-JAN-2003 (first entry)  
XX  
DE SVBV type 2 promoter useful in RNAi vector for nematode control.  
XX  
XX RNAi molecule; double-stranded interfering RNA; nematode control;  
KW RNA mediated interference; mRNA transcript; nematode gene; growth;  
KW development; parasitism; reproduction; RNAi vector; mRNA translation;  
KW nematode inhibitor; agricultural industry; anti-nematode; ds.  
XX  
OS Strawberry vein banding virus; type 2.  
XX  
PN WO200196584-A2.  
XX  
PD 20-DEC-2001.  
XX  
XX 12-JUN-2001; 2001WO-US018911.  
XX  
PR 12-JUN-2000; 2000US-0210917P.  
XX  
XX (AKKA-) AKKADIX CORP.  
XX  
XX Mushegian AR, Taylor CG, Feitelson JS, Broshkin AM;  
XX WPI; 2002-139714/18.  
XX  
XX RNA mediated interference molecule useful for disrupting cellular process  
PT in a nematode, for controlling nematodes comprises genetic regulatory  
PT sequences.  
XX  
XX Example 5; Page 59; 103pp; English.  
XX  
XX The present invention relates to RNAi (double-stranded interfering RNA or  
CC RNA mediated interference) molecules (nematode genes), and methods of  
CC using these sequences in nematode control. RNAi molecules selectively  
CC target mRNA transcripts of essential nematode genes. The RNAi molecules  
CC of the invention are useful for disrupting cellular processes in a  
CC nematode by contacting the nematode with a composition comprising an RNAi  
CC molecule. The RNAi molecules are useful for killing nematodes and/or  
CC inhibiting their growth, development, parasitism or reproduction and also  
CC for the regulation of levels of specific mRNA in nematodes. Genetic  
CC regulatory sequences such as promoters, enhancers and terminators can be  
CC used in genetic constructs such as RNAi vectors which can be used for  
CC nematode control. The RNAi molecules are capable of targeting and  
CC reducing (and, in some cases, preventing) the translation of a specific  
CC gene product, and can be used to reduce or prevent mRNA translation in  
CC any tissue of the nematode because of its ability to cross tissue and  
CC cellular boundaries. The RNAi molecule can be contacted with a nematode  
CC by soaking, injection, or consumption of a food source containing an RNAi  
CC molecule. The RNAi molecules can also be used as an epigenetic factor to  
CC prevent the proliferation of subsequent generations of nematodes, to  
CC produce nematode inhibitors or RNAi in the plants, and provide new  
CC biotechnological strategies for managing nematodes under sustainable  
CC agricultural conditions. ABX09824-ABX09833 represent exemplary genes  
XX useful in RNAi vectors for the control of nematodes. (Updated on 29-AUG-  
XX 2003 to standardise OS field)  
XX  
SQ Sequence 1071 BP; 372 A; 232 C; 202 G; 265 T; 0 U; 0 Other;

Query Match 100.0%; Score 476; DB 6; Length 1071;  
Best Local Similarity 100.0%; Pred. No. 2.4e-123;  
Matches 476; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AACTATGCTGATGACAGATTAATCTTAATAGCAATTAATCAGAAATTAATCAGGAGAAA 60  
DB 563 AACTATGCTGATGACAGATTAATCTTAATAGCAATTAATCAGAAATTAATCAGGAGAAA 622  
QY 61 GAATTAATAACTCTTTTCAAGATATGAAGCCGCTTTTCAAGTGCCAGCTAGCTATCACT 120  
DB 623 GAATTAATAACTCTTTTCAAGATATGAAGCCGCTTTTCAAGTGCCAGCTAGCTATCACT 682  
QY 121 GAAAGACAGCAAGCAATGTGCTGATGACACAGAAACACATCTTTTGACGAGATGT 180  
DB 683 GAAAGACAGCAAGCAATGTGCTGATGACACAGAAACACATCTTTTGACGAGATGT 742  
QY 181 GAAGCAGCCAGAGTGGTCCCAAGACGCACTCAGAAAAGGCATCTTACCGACACAGAA 240  
DB 743 GAAGCAGCCAGAGTGGTCCCAAGACGCACTCAGAAAAGGCATCTTACCGACACAGAA 802  
QY 241 AAAGCAACCCAGCTCATCTCAACATGTAGACTGTGCTTATGCGTTCGCTGAAGATA 300  
DB 803 AAAGCAACCCAGCTCATCTCAACATGTAGACTGTGCTTATGCGTTCGCTGAAGATA 862  
QY 301 AGACTGACCCAGCCAGCACTTAAGAGAAATTAAGTGGTCTAGCTCCACTTTA 360  
DB 863 AGACTGACCCAGCCAGCACTTAAGAGAAATTAAGTGGTCTAGCTCCACTTTA 922  
QY 361 GCTTTAATAATATGTTTCATTAATTAATCTCTGCTTTTCTCTCTATATAAGAGCTTGT 420  
DB 923 GCTTTAATAATATGTTTCATTAATTAATCTCTGCTTTTCTCTCTATATAAGAGCTTGT 982  
QY 421 ATTTCATTTAGGCGAGGCGAACAACACACAGAAACCTCCCTGCTTACAAACC 476  
DB 983 ATTTCATTTAGGCGAGGCGAACAACACACAGAAACCTCCCTGCTTACAAACC 1038

RESULT 5  
AAD39134  
ID AAD39134 standard; DNA; 651 BP.  
XX AC AAD39134;  
XX DT 04-OCT-2002 (first entry)  
XX DE Strawberry vein banding virus (SVBV) promoter DNA, sesVBV.  
XX DW Strawberry vein banding virus promoter; transgenic plant; transgenic; ds.  
XX OS Strawberry vein banding virus.  
XX PN WO200240691-A2.  
XX PD 23-MAY-2002.  
XX PF 30-OCT-2001; 2001WO-US047964.  
XX PR 01-NOV-2000; 2000US-0245354P.  
XX PA (MAXY-) MAXYGEN INC.  
XX PI Wu G, McBride K;  
XX DR WPI; 2002-479908/51.  
XX PT Novel isolated or recombinant nucleic acid comprising strawberry vein  
XX PT banding virus promoter operably linked to a heterologous polynucleotide,  
XX PT useful for expressing the heterologous polynucleotide in a plant cell.  
XX PS Claim 1; Page 47-48; 48pp; English.  
XX CC The invention relates to an isolated/recombinant nucleic acid comprising  
XX CC a strawberry vein banding virus (SVBV) promoter operably linked to a

CC heterologous polynucleotide. The invention is useful for expressing a  
CC heterologous polynucleotide in a plant cell, by introducing the nucleic  
CC acid into a plant cell, where the plant cell is present within a plant  
CC and Agrobacterium is used to introduce the isolated nucleic acid into the  
CC cell. The invention is useful to drive gene expression in plant cells and  
CC transgenic plants, for transforming plant cells and producing transgenic  
CC plants, for compensating an endogenous protein at higher than normal levels,  
CC plant, for expressing a novel gene in a plant, and to suppress expression of  
CC endogenous plant genes. The present sequence is sesVBV promoter DNA  
XX  
SQ Sequence 651 BP; 213 A; 167 C; 132 G; 139 T; 0 U; 0 Other;

Query Match 77.5%; Score 369; DB 6; Length 651;  
Best Local Similarity 100.0%; Pred. No. 2.2e-93;  
Matches 369; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 108 GCTAGCTATCACTGAAAACAGACAGCAAGCAATGCTGCTCGATGCACAGACCACTCT 167  
DB 278 GCTAGCTATCACTGAAAACAGACAGCAATGCTGCTCGATGCACAGACCACTCT 337  
QY 168 TTGCAGCAGATGTGAAGCAGCCAGAGTGTGTCACCAAGACGCACTCAGAAAAGGCATCTTC 227  
DB 338 TTGCAGCAGATGTGAAGCAGCCAGAGTGTGTCACCAAGACGCACTCAGAAAAGGCATCTTC 397  
QY 228 TACCGACAGAAAAGACCAACAGCTCATCTCAACATGTAGACTGTGCTTATGGG 287  
DB 398 TACCGACAGAAAAGACCAACAGCTCATCTCAACATGTAGACTGTGCTTATGGG 457  
QY 288 TCGGCTGAAGATAAGACTGACCCAGCCAGCACTTAAAGAGAAAATAATGCAAGTGGTCC 347  
DB 458 TCGGCTGAAGATAAGACTGACCCAGCCAGCACTTAAAGAGAAAATAATGCAAGTGGTCC 517  
QY 348 TAGCTCCACTTTAGCTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 407  
DB 518 TAGCTCCACTTTAGCTTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 577  
QY 408 ATAAAGAGCTTCTATTTTTCATTTGAAGCGAGGCGAACAACACACAGAAACCTCCCTGC 467  
DB 578 ATAAAGAGCTTCTATTTTTCATTTGAAGCGAGGCGAACAACACACAGAAACCTCCCTGC 637  
QY 468 TTACAAACC 476  
DB 638 TTACAAACC 646

RESULT 6  
AAD39135  
ID AAD39135 standard; DNA; 701 BP.  
XX AC AAD39135;  
XX DT 04-OCT-2002 (first entry)  
XX DE Strawberry vein banding virus (SVBV) promoter DNA, lesVBV.  
XX DW Strawberry vein banding virus promoter; transgenic plant; transgenic; ds.  
XX OS Strawberry vein banding virus.  
XX PN WO200240691-A2.  
XX PD 23-MAY-2002.  
XX PF 30-OCT-2001; 2001WO-US047964.  
XX PR 01-NOV-2000; 2000US-0245354P.  
XX PA (MAXY-) MAXYGEN INC.  
XX PI Wu G, McBride K;  
XX DR WPI; 2002-479908/51.

XX	Novel isolated or recombinant nucleic acid comprising strawberry vein
PT	banding virus promoter operably linked to a heterologous polynucleotide,
PT	useful for expressing the heterologous polynucleotide in a plant cell.
XX	Claim 1; Page 48; 48pp; English.
XX	The invention relates to an isolated/recombinant nucleic acid comprising
CC	a strawberry vein banding virus (SVBV) promoter operably linked to a
CC	heterologous polynucleotide. The invention is useful for expressing a
CC	heterologous polynucleotide in a plant cell, by introducing the nucleic
CC	acid into a plant cell, where the plant cell is present within a plant
CC	and Agrobacterium is used to introduce the isolated nucleic acid into the
CC	cell. The invention is useful to drive gene expression in plant cells and
CC	transgenic plants, for transforming plant cells and producing transgenic
CC	plants, for compensating for missing or altered gene expression in a
CC	plant, for expressing an endogenous protein at higher than normal levels,
CC	for expressing a novel gene in a plant, and to suppress expression of
CC	endogenous plant genes. The present sequence is lesVBV promoter DNA
XX	
SQ	Sequence 701 BP; 222 A; 176 C; 139 G; 164 T; 0 U; 0 Other;
	Query Match 77.5%; Score 369; DB 6; Length 701;
	Best Local Similarity 100.0%; Pred. No. 2.2e-93;
	Matches 369; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	108 GCTAGCTATCACTGAAAGACAGCAAGCAATGGTGTCGATGCACCAAGACCATCT 167
DB	319 GCTAGCTATCACTGAAAGACAGCAAGCAATGGTGTCGATGCACCAAGACCATCT 378
QY	168 TTTCGACGAGATGTGAAGCAGCAGAGTGGTCCACAGCGCATCTCAGAAAAGGCTATCTTC 227
DB	379 TTTCGACGAGATGTGAAGCAGCAGAGTGGTCCACAGCGCATCTCAGAAAAGGCTATCTTC 438
QY	228 TACCCGACAGAAAAGACACACCACAGCTCATCTCCAAACATCTAGACTGTGTTATGGC 287
DB	439 TACCCGACAGAAAAGACACACCACAGCTCATCTCCAAACATCTAGACTGTGTTATGGC 498
QY	288 TCGGCTGAAGATAAGACTGACCCGAGGCGACACTAAAGAAGAAATAATGCAAGTGGTCC 347
DB	499 TCGGCTGAAGATAAGACTGACCCGAGGCGACACTAAAGAAGAAATAATGCAAGTGGTCC 558
QY	348 TAGCTCCACTTTAGCTTTAAATAATTATGTTTCATTATTATCTCTGTTTGTCTCTAT 407
DB	559 TAGCTCCACTTTAGCTTTAAATAATTATGTTTCATTATTATCTCTGTTTGTCTCTAT 618
QY	408 ATAAAGAGCTTGTTATTTTCATTGAAGCGACAGGCGCAACACACACAGAACCTCCCTGC 467
DB	619 ATAAAGAGCTTGTTATTTTCATTGAAGCGACAGGCGCAACACACACAGAACCTCCCTGC 678
QY	468 TTACAAACC 476
DB	679 TTACAAACC 687
RESULT 7	
ABL26902/c	
ID	ABL26902 standard; DNA; 3294 BP.
XX	ABL26902;
AC	
XX	
XX	26-MAR-2002 (first entry)
DT	
XX	Drosophila melanogaster genomic polynucleotide SEQ ID NO 32179.
DE	
XX	Drosophila; developmental biology; cell signalling; insecticide;
XX	pharmaceutical; gene; ds.
KW	
XX	Drosophila melanogaster.
OS	
XX	
FN	WO200171042-A2.
XX	
PD	27-SEP-2001.

[illegible]

```
PA (CHIR ) CHIRON CORP.
XX
PI Frazer CM, Hickey E, Peterson J, Tettelin H, Venter JC, Scarlato V;
PI Masignani V, Galeotti C, Mora M, Ratti G, Scarselli M,
PI Rappuoli R, Pizza M;
XX
XX WPI; 2000-318079/27.
XX
XX Isolated nucleotide sequences of Neisseria meningitidis which can be used
PT in the diagnosis and treatment of N. meningitidis infection and other
PT Neisserial infections, for example, N.gonorrhoea.
XX
PS Claim 7; Page 330-353; 1760pp; English.
XX
XX The present invention describes methods of obtaining immunogenic proteins
CC from Neisseria genomic sequences. AAA81453 to AAA82414 represent
CC specifically claimed Neisseria meningitidis genomic DNA sequences;
CC AAA81260 to AAA81303 and AAA82562 to AAA82563 represent Neisseria DNA
CC sequences and their corresponding proteins; AAA81254 to AAA81259 and
CC AAA81304 to AAA81321 represent PCR primers used in the isolation of
CC Neisseria meningitidis DNA sequences; and AAA81322 to AAA81452 represent
CC Neisseria meningitidis MenB polynucleotide ORF sequences, which are all
CC used in the exemplification of the present invention. The nucleic acid
CC sequences, protein sequences, and antibodies against them, can be used in
CC the manufacture of a composition. The composition can be used as a
CC medicament (or in the manufacture of a medicament) for treating,
CC preventing or diagnosing infection due to Neisserial bacteria. For
CC example, some of the identified proteins could be components of vaccines
CC against Meningococcus B; against all serotypes; and/or against all
CC pathogenic Neisseriae. Identification of sequences from the bacterium
CC will also facilitate production of biological probes, particularly
CC organism-specific probes. Attempts to make efficacious Meningococcus B
CC vaccines have failed mainly due to antigen tolerance. Multivalent
CC vaccines have also been tried but none have successfully overcome
CC antigenic variability. The provision of further, complete sequences may
CC provide an opportunity to identify secreted or surface exposed proteins
CC that may be presumed targets for the immune system and which are not
CC antigenically variable or at least more conserved than other more
CC variable regions
XX
SQ Sequence 78845 BP; 19372 A; 19372 C; 19042 G; 19318 T; 0 U; 2 Other;

Query Match
Best Local Similarity 7.9%; Score 37.6; DB 3; Length 78845;
Matches 106; Conservative 0; Mismatches 114; Indels 0; Gaps 0;

QY 217 AAGGATCTTCTACCGACACAGAAACACACACAGCTCATCATCAATGTAGACT 276
DB 11186 AAGGATCTTCTACCGACACAGAAAGAAATAGAGCCGAGGATTTGATGGATGTTTACC 11245
QY 277 GTCGTTATGCTGCGCTGAAGATAAGACTGACCCAGCCGACCTAAAGAAAGAAATAT 336
DB 11246 GTGCTTTGATTCGGTTGGAATATTCGAAATTCGATTTCCGATAGTAAAGACAAAAG 11305
QY 337 GCAAGTGGTCTAGCTCCACTTTAGCTTTAATTAATTTGTTTCAATTTCTCTGCTT 396
DB 11306 GGTTTTGGTAAGTACCAAAATCAAAACAGTTGATTTGATTCGGCAGATATGTTCTCTG 11365
QY 397 TTGCTCTCTATATAAGAGCTTGTATTTTCATTTGAAGGC 436
DB 11366 ATGATCTTTTCTGAAATGCTTACAGCTTCATATAGGC 11405

RESULT 9
AAA81490 08/C
Continuation (9 of 15) of AAA81490 from base 800001 (N. meningitidis B full length genome)
WP Sequence split into 15 fragments LOCUS AAA81490 Accession AAA81490
WP Fragment Name Begin End
WP AAA81490_00 1 110000
WP AAA81490_01 100001 210000
WP AAA81490_02 200001 310000
WP AAA81490_03 300001 410000
WP AAA81490_04 400001 510000

PA Pizza M, Hickey E, Peterson J, Tettelin H, Venter JC;
PI Masignani V, Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V;
PI Rappuoli R, Frazer CM, Grandi G;
XX
XX WPI; 2000-647603/62.
XX
XX Neisseria meningitidis B full length genome sequence and open reading
PT frames are used to detect, treat and prevent Neisserial infections.
XX
PS Claim 7; Appendix A; 692pp; English.
XX
CC The present invention describes the full length genome of Neisseria
```







```

RESULT 15
AAD53224_3
Continuation (4 of 6) of AAD53224 from base 300001 (Human chromosome 3 q-arm breakpoint
WP Sequence split into 6 fragments LOCUS AAD53224 Accession AAD53224
WP Fragment Name Begin End
AAD53224_0 1 110000
AAD53224_1 100001 210000
AAD53224_2 200001 310000
AAD53224_3 300001 410000
AAD53224_4 400001 510000
AAD53224_5 500001 567571

Query Match 7.6%; Score 36.2; DB 7; Length 110000;
Best Local Similarity 56.2%; Pred. No. 38;
Matches 68; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY 317 AGCACTAAGAGAGAAATATGCGAGTGGTCCCTAGCTCCACTTTAGCTTTAATAATATGT 376
Db 3031 AGCACAGGCAATTTTAAAGTGAACCTGCTTCITGCTCATTTGCACATATAATTTTGA 3090

QY 377 TTCATTATTATCTCTGCTTTTCTCTCTATATAAAGAGCTTGATTTTCATTGAAGGC 436
Db 3091 GTCATCTTATTGACTTCITTTGAACCTCTATCCCAACCTGAGGGCAGATTAGAGGG 3150

QY 437 A 437
Db 3151 A 3151

```

Search completed: July 31, 2004, 08:13:44  
Job time : 313 secs

